STEREOSELECTIVE SYNTHESES AND SELF-ASSEMBLING CAPABILITIES OF HETEROCYCLIC CINNAMIC ACIDS

PhD Dissertation

Csankó Krisztián

Supervisor: Dr. Pálinkó István

Co-supervisor: Dr. Sipos Pál

Doctoral School of Chemistry

Department of Organic Chemistry

Faculty of Science and Informatics

University of Szeged

Szeged, Hungary

2015

This dissertation is available online at Library of University of Szeged:
http://doktori.bibl.u-szeged.hu/cgi/search/advanced

Recommended Citation:

K. Csankó, "Stereoselective syntheses and self-assembling capabilities of heterocyclic cinnamic acids", PhD Thesis, University of Szeged, 2014.

TABLE OF CONTENTS

.....I

List of abbreviations.....VII

1. Introduction 1

2. Literature Overview.....2

2.1 Brief history of cinnamic acid.....2

2.1.1 The relation of cinnamic acid and cinnamon.....2

2.1.2 Discovery and the first synthesis of cinnamic acid.....3

2.1.3 First applications of cinnamic acid and its derivatives.....4

2.2 Natural occurrence and the synthesis of cinnamic acid and its derivatives 6

2.3 Synthetic ways to E-cinnamic acids 6

2.3.1 The Perkin condensation reaction.....6

2.3.2 The Claisen-Schmidt condensation reaction.....7

2.3.3 The Knövenagel-Döbner condensation reaction.....7

2.3.4 The Heck coupling reaction 8

2.3.5 Microwave-assisted synthesis of cinnamic acid derivatives 9

2.4 Synthetic ways to *Z*-cinnamic acids.....9

2.4.1 Photoinduced isomerization of E-cinnamic acids.....9

2.4.2 Z-selective catalytic hydrogenation of alkynes.....10

2.4.3 The Horner-Wadsworth-Emmons (HWE) olefination.....11

2.4.4 Z-selective olefin metathesis.....13

2.5 Synthetic ways to α-substituted cinnamic acids 13

2.5.1 The modified Perkin and other relevant condensations.....13

2.5.2 Synthesis of α-substituted cinnamic acids via addition reactions.....14

2.6 Secondary chemical bonds and the self-assembling capabilities of cinnamic acids 15

2.6.1 About secondary chemical bonds 15

2.6.2 Possible interactions between cinnamic acids 16

2.6.3 Self-assembling of various molecules over transition metal surfaces 19

3. The main aims of the dissertation 22

4. Experimental part 23

4.1 Materials 23

4.1.1 The molecules studied 23

4.1.2 Solvents for NMR spectroscopy 24

4.1.3 Polycrystalline metal surfaces 24

4.2 Instrumental methods 25

4.2.1 Analytical and structure analysis methods 25

4.2.2 Surface inspection techniques 27

4.2.3 Reactors and preparative instruments 28

4.3 General methods for the stereoselective synthesis of *E*-cinnamic acids 29

4.3.1 Preparation of 2-selenophenylcarboxaldehyde with Vilsmeier-Haack formylation 29

4.3.2 Preparation of E-cinnamic acid with the Perkin condensation 30

4.3.3 Preparation of E-ethylcinnamate with the Claisen-Schmidt condensation 30

4.3.4 Preparation of diethyl-2-benzylidene malonate with the Knövenagel condensation 31

4.3.5 Microwave-assisted Knövenagel-Döbner condensation procedures 31

4.3.6 Preparation of E-cinnamic acid with the Knövenagel-Döbner condensation 32

4.3.7 Preparation of E-3-(2 or 3-furyl-, 2 or 3-thienyl-, 2, 3 or 4-pyridyl- and 2-selenophenyl)propenoic acids with the Knövenagel-Döbner condensation 32

4.3.8 Solubilities of cinnamic acid derivatives in DMSO 33

4.3.9 Preparation of octyl-E-3-(2-thienyl)propenoate and octyl-E-3-(3-thienyl)propenoate....34
4.4 General methods for the synthesis of *E*-*Z*-cinnamic acid isomer mixtures....35
4.4.1 Preparation of E-Z cinnamic acid isomer mixtures via UV induced photoisomeriza-tion and the kinetics of the process....35
4.5 General methods for the stereoselective synthesis of *Z*-cinnamic acids....36
4.5.1 Preparation of 3-phenylpropiolic acid and methyl 3-phenylpropiolate....36
4.5.2 Z-selective hydrogenation of 3-phenylpropiolic acid with Lindlar's catalyst 37
4.5.3 Utilizing the HWE synthesis for Z-selective preparation....39
4.5.4 Preparation of the different HWE phosphonates....39
4.5.5 Optimization of the Z-selective HWE reaction with Still-Gennari and Ando-type phosphonates....41
4.6 Synthesis of the α-substituted cinnamic acid derivatives....44
4.6.1 The synthesis of E-Z-3-phenyl-2-(3-pyridyl)propenoic acid via the modified Perkin condensation....45
4.6.2 The synthesis of E-2-phenyl-3-(3-pyridyl)propenoic acid via the HWE method 46
4.7 Computational chemistry part: software and methods....47
4.7.1 Official software packages....47
4.7.2 Creating software for conformational analysis (Quantum Hyppo and Izzy Reader) 47
4.7.3 The applied quantum chemical methods....51
5. Results and Discussion....54
5.1 Summary of the synthetic part....54
5.1.1 The results of the E-selective methods....54
5.1.2 Photoisomerization of E-cinnamic acids....56

5.1.3 Summary of the Z-selective methods....57

5.1.4 Summary of the synthetic methods for α-substituted cinnamic acid derivatives....58

5.2 Determining the different hydrogen bonds with infrared spectroscopy....58

5.3 $^{1}H-^{1}H$ correlation (COSY) and nuclear Overhauser effect (NOESY) NMR spectra of heteroaryl cinnamic acids....64

5.4 Identifying the π-stacking interactions with powder X-ray diffractometry....75

5.5 Self-assembling of the studied molecules over polycrystalline gold and silver surfaces....77

5.6 Computational results....83

5.6.1 Conformational analysis....83

5.6.2 Hydrogen-bonded dimers....87

5.6.3 Hydrogen-bonded tetramers and other structures....91

5.6.1 Π-stacking interactions....94

Summary....96

Related Publications....A

Other Publications....B

References....C

Magyar nyelvű összefoglaló (Hungarian summary)....I

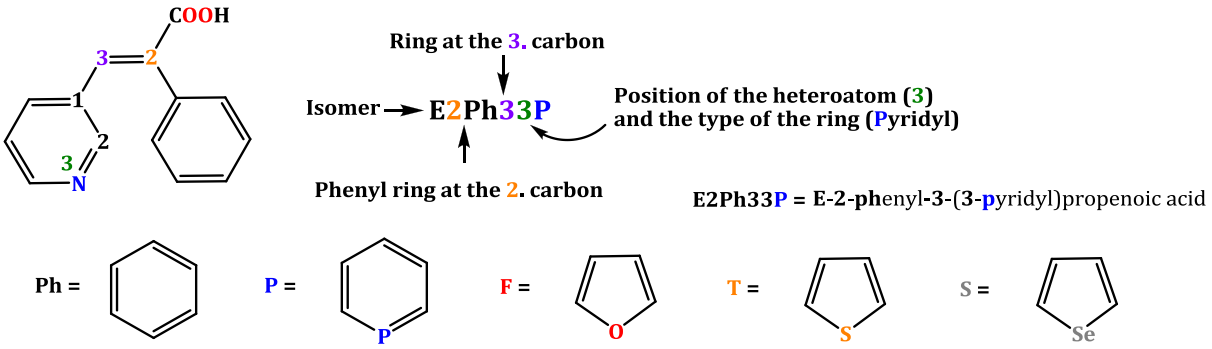
ACKNOWLEDGEMENT....L

Appendix....M

LIST OF ABBREVIATIONS

MATERIALS

General method for the naming of the studied cinnamic acids:



A few examples for the shortened names:

E3Ph – cinnamic acid (E-3-**ph**enylpropenoic acid (**1/E**))

E32F – E-3-(2-**f**uryl)propenoic acid

E32T – E-3-(2-**t**hienyl)propenoic acid

E32S – E-3-(2-**s**elenienyl)propenoic acid

E33P – E-3-(3-**p**yridyl)propenoic acid

E23T33T – E-2-(3-**t**hienyl)-3-(3-**t**hienyl)propenoic acid

Z2Ph33P – Z-2-**p**henyl-3-(3-**p**yridyl)propenoic acid

Z – "Z" instead of "E" before the shorthand names of the materials listed above, means the *Z* isomer

M – at the end of the shorthand names of the materials listed above, means the individual monomer

D – at the end of the shorthand names of the materials listed above, means the dimer kept together by double *hydrogen bonds* between the *carboxylic* groups

T – at the end of the shorthand names of the materials listed above, means the tetramer built up from two dimers via (aromatic)C–H…X *hydrogen bond* between the aromatic rings

Me – at the end of the shorthand names of the materials listed above, means the *methyl ester* of the corresponding cinnamic acid derivative

O – at the end of the shorthand names of the materials listed above, means the octylester of the corresponding cinnamic acid derivative

ZW – at the end of the shorthand names of the materials listed above, means that the pyridyl derivatives are in the zwitterionic forms

EWG – electron withdrawing group

Ar – aromatic group

Ph – phenyl group

R – substituent

X – heteroatom (O,N,S)

Hlg – halogen (F, Cl, Br, I)

Ac – **a**cetyl group

Me – methyl group

Et – ethyl group

Pr – propyl group

MeO – methoxy group

EtO – ethoxy group

TEA – triethylamine

THF – tetrahydrofuran

DMSO – dimethyl sulfoxide

DMF – *N,N*-dimethylformamide

MeOH – methanol

EtOH – ethanol

AcOH – acetic acid

EtOAc – ethyl acetate

Ac_2O – acetic anhydride

KHMDS – potassium bis(trimethylsilyl)amide

***t*-BuOK** – potassium tert-butoxide

DMAP – 4-(dimethylamino)pyridine

PPE – polyphosphate ester

TBAB – tetrabutylammonium bromide

TMS – tetramethylsilane

TMSO – trimethylsilyloxy group

DSS – 4,4-dimethyl-4-silapentane-1-sulphonic acid

METHODS

MALDI-MS – matrix-assisted laser desorption/ionization mass spectrometry

(M)IR – mid-range infrared spectroscopy

MCT – mercury cadmium telluride detector

DTGS – deuterated triglycine sulfate

ATR–IR – attenuated total reflection infrared spectroscopy

UV–Vis – ultraviolet–visible spectroscopy

NMR – nuclear magnetic resonance spectroscopy

COSY – nuclear magnetic resonance correlation spectroscopy

NOESY – nuclear Overhauser effect spectroscopy

HSQC – heteronuclear single quantum coherence spectroscopy

XRD – X-ray diffractometry

IRM – infrared microspectroscopy

SEM – scanning electron microscopy

AFM – atomic force microscopy

GC – gas chromatography

FID – flame ionization detector

HPLC – high pressure liquid chromatography

PLD – pulsed laser deposition

MW – microwave irradiation

UV – ultraviolet irradiation

hν – irradiation with light

US – ultrasound irradiation

SAM – self-assembled monolayer

LbL – layer-by-layer

HWE – Horner-Wadsworth-Emmons

OBF – one bond flip

ROCM – ring opening cross-metathesis

UWFP – used without further purification

OTHER

*– indicates data from own measurements

rbf. – round bottomed flask

cat. – catalytic amount/catalyst

sicc. – water-free compound

cc. – concentrated

r.t. – room temperature

mp. – melting point

Mr. – relative molecular mass

R_f – retention factor

eq. – equivalent

quant. – quantitative

M – concentration in mol/dm^3 (c)

K – equilibrium constant

k_f – forward rate constant

k_b – backward rate constant

ε – molar extinction coefficient (M^{-1} cm^{-1})

λ – wavelength

Δ – heating/elevated temperature

MCU – microcontroller unit

AVR – advanced virtual RISC (Reduced Instruction Set Computing) by Atmel co.

ISP – in system programmer

USB – universal serial bus

MOSFET – metal oxide semiconductor field-effect transistor

CPU – central processing unit (processor)

E – energy of the optimized molecular framework (Joule, kJoule/mol, kcal/mol)

BSSE – basis set superposition error

RMS – root mean squared gradient

CP – counterpoise method

He – energy in Hartree

D° – dihedral angle

ZPE – zero-point energy

HF – Hartree-Fock level

DFT – density functional theory

B3LYP – Becke three-parameter Lee-Yang-Parr method

CBSB7 – covers the 6-311G(2d,d,p) basis set

PM(3,6) – parameterized model number 3 or 6 method

MP2 – Møller-Plesset's second perturbation theory

WB97XD – Head-Gordon type Becke function with empirical dispersion

ABBREVIATIONS OF THE NMR SPECTRA

υ – frequency (Hz)

J – coupling constant (Hz)

δ – chemical shift (ppm)

ppm – parts per million

s – singlet

d – doublet

t – triplet

q – quartet

quint. – quintet

m – multiplet

dd –doublet of doublets

dt – doublet of triplets

td – triplet of doublets

br – broad signal

„Some scientists claim that hydrogen, because it is so plentiful, is the basic building block of the universe. I dispute that. I say that stupidity is far more abundant than hydrogen, and that is the basic building block of the universe."

Frank Zappa

1. INTRODUCTION

Cinnamic acid (*E-3-phenylpropenoic acid*, **1/E**) is a well-known material, because it can be found in *cinnamon*, which was mentioned for the first time on ancient Egyptian hieroglyphs (2000 B.C)[2]. The oil of leafs and the bark of *Cinnamomum Verum* (cinnamon plants are native in *Sri Lanka*) contains *cinnamaldehyde* (**2**) as a main component (50–80%) and several percent of *esters* and *terpenes.*[3] *Cinnamon* (the bark of the plant) and the oil of the leafs have been widely used as flavor and food preservative[4] as well for thousands of years, because some of its components, mostly the *cinnamic acid* derivatives, have antimicrobial effect on pathogenic bacteria[5] and on human diseases like upper respiratory tract infection caused by *Pneumococcus.* As it was mentioned, *cinnamic acid* can be found in *cinnamon*, due to the slow oxidation of *cinnamaldehyde* (**2**) to the more stable *E-cinnamic acid* (**1/E**) on contacting with air.[6]

Cinnamic acid (**1/E**) was synthesized first by W.H. Perkin[7] in 1868 but, its use and those of its derivatives were delayed until the second world war, where the first UV-protective sunscreen was prepared from *cinnamate esters*[8] for the soldiers fighting in the Pacific area. The second "big boom" in the history of *cinnamic acid*, came at the late 80's, when they gained importance as sample matrices in the MALDI–MS technique[9] due to its high absorbance in the UV range (the average molar extinction coefficient is around: 19556 $M^{-1}cm^{-1}$).

The substituted derivatives of *cinnamic acid*s and some heteroatom-containing derivatives came into the focus of our research group, since they proved to be good models for studying *hydrogen bonding* driven assembling in solutions as well as in the solid state.[10,11,12,13,14] It was found that they were capable short- and long-range ordering in solution and the solid state, respectively. The fundamental units were found to be the dimers, kept together by strong *hydrogen bonds*, while the multimers of the dimers interacting with weaker (aromatic)C–H…X (where X = O, N, S) close contacts.

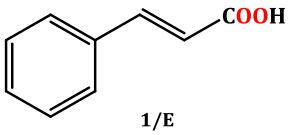
In most of these studies, the combination of spectroscopic methods and molecular modelling was applied for the characterization of the structure-forming interactions.

2. LITERATURE OVERVIEW

2.1 Brief history of cinnamic acid

2.1.1 The relation of cinnamic acid and cinnamon

Cinnamic acid (*Figure 1*) **1/E** (and its derivatives) is one of the several substances that we widely use in our everyday life without the recognition of its presence.



1/E

Figure 1 The structure of *E-cinnamic acid* (*E-3-phenylpropenoic acid*, **1/E**)

Cinnamic acid, or *E-3-phenylpropenoic acid* **1/E**, is a white crystalline material with honey-like odor. It has a completely planar structure, due to the continuous conjugation along the molecular framework.[15] It appears in many plants over the world (e.g., *Liquidambar orientalis*[16], *Vitellaria paradoxa*[17], *Petroselinum tuberosum*[18]), but the most important sources are the various types of *cinnamons*. *Cinnamon* is one of the oldest spices we use. It has very pleasant smell and taste, and has beneficial effect on digestion and health, too. There are many different theories about the first mention of *cinnamon* in written history, because of the confusion around the four general types of *cinnamon* plants in the *Lauraceae* family[19], which has worldwide distribution in tropical and warm climates:

- *Cassia, or "wild cinnamon"* the bark of *Cinnamomum Iners* tree, can be found in *Ethiopia* and *Singapore*[19],
- *Malabathrum,* the grounded leafs of *Cinnamomum Tamala*, which is native in *India* and *China*[20],
- *Serichatum,* the aromatic bark of *Cinnamomum Cassia,* widely cultivated in *China*, *Laos*, *Malaysia* and *Vietnam*[20],
- *And the "true cinnamon"* is the inner bark of the *Cinnamomum Verum* plant (*Figure 2*).

Figure 2 *The Cinnamomum verum* plant[1]

The oil of the *Cinnamomum verum* bark contains many compounds, such as *cinnamaldehyde* **2** in more than 50%, *E-ethyl cinnamate* **3**, *R-(-)-linalool* **4**, *eugenol* **5,** *estragole* **6** and several *terpenes*[3] (*Figure 3*).

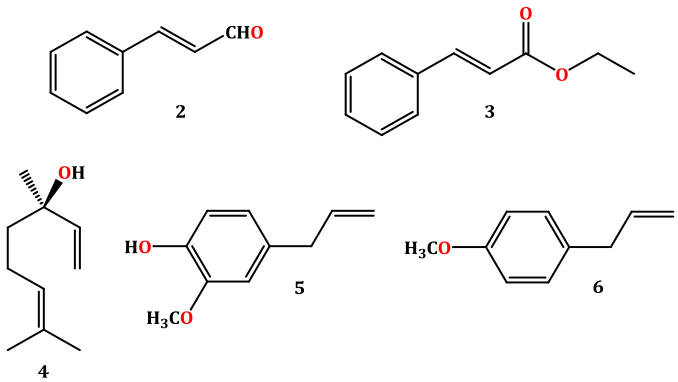


Figure 3 Main components of *cinnamon* essential oil

2.1.2 Discovery and the first synthesis of cinnamic acid

Despite the high concentration of *cinnamaldehyde* **2** in the essential oil, it was isolated quite late, in 1834 by *Dumas* and *Préligot*[21]; then, it was first synthesized in 1854 by *Chiozza*[22], via the *aldol* condensation of *benzaldehyde* **7** and *acetaldehyde*. *Cinnamaldehyde* **2** is slowly oxidized[6] to *cinnamic acid* **1/E** on contact with air; thus, the essential oil of *cinnamon* will contain significant amount of *cinnamic acid* **1/E** after ageing. In spite of this, *cinnamic acid* **1/E** was not discovered through isolation from *cinnamon* oils. It was first synthesized in 1868 by *Perkin*[7], who discovered the

condensation reaction between *acetic anhydride* **8** and *aromatic aldehydes*, such as *salicylaldehyde* **9** giving *o-coumaric acid* **10** that can be transformed to *coumarine* **11** (via intramolecular esterification) or *benzaldehyde* **7** giving *cinnamic acid* **1/E** in the presence of the *potassium*, or *sodium salt* of the corresponding acid (*acetic acid*) (*Figure 4*). However, the importance of *cinnamic acid* **1/E** was smaller than that of *coumarine* **11** in those days. *Coumarine* was an important substance in the perfume industry, it was the first synthetic perfume in history.[7]

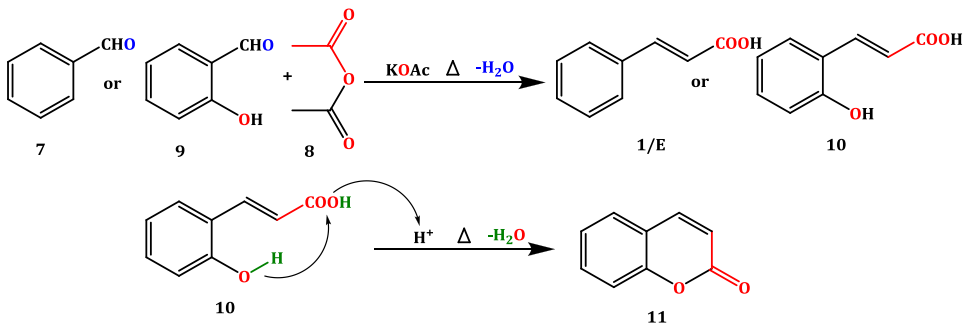


Figure 4 The *Perkin* condensation reactions

2.1.3 First applications of cinnamic acid and its derivatives

The first mention about the application of *cinnamic acid* itself, was delayed to 1946, when the Swiss chemist *Franz Greiter* introduced the first modern sunscreen called *"Gletscher Créme"*[8] that contained *etylhexylmetoxycinnamate* **12** (*Figure 5*).

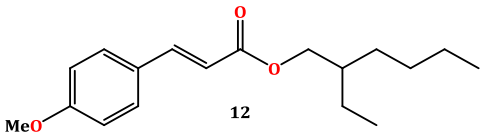


Figure 5 The structure of *(RS)-2-ethylhexyl (2E)-3-(4-methoxyphenyl)propenoate* (**12**)

It is the main component (10–20%) of sunscreens up to now, due to its wide absorption band in the UV B/C range (260–315nm)[23] and the high molar absorption coefficient (ε= *19556 $M^{-1}cm^{-1}$ for *E-cinnamic acid* **1/E** and *24160 $M^{-1}cm^{-1}$ for *E-ethyl cinnamate* **3**). After the "big boom" on the market of sunscreens, the developers needed a more efficient way to synthesize *cinnamic acids* than the old *Perkin* condensation, which performs well (yield 85-93%) in the laboratory, but the scalability is poor. In 1951 J. *Blake et al.*[24] published a very general method, based on the *Knövenagel-Döbner*

condensation[7], which can be used for the quick and efficient synthesis of a variety of *cinnamic acid* derivatives.

In the second half of the 20th century, the direct use of *cinnamic acid* in cosmetics started to grow very fast, and it became a basic material in the production of fine fragrances, decorative cosmetics, shampoos, toilet soaps or even in household chemicals, such as cleaners and other detergents.[25]

The next evolutionary step for *cinnamic acids* was the invention of the *matrix-assisted laser desorption mass spectrometry* (MALDI–MS) in 1985 by *Franz Hillenkamp*[26] and his postdoc. They found that *amino acids*, even *proteins* and other large molecules can be ionized more easily with UV light, when they were mixed with conjugated aromatic compounds like *cinnamic acid* derivatives, because then, they had a wide absorption band in the UV range.[27,28,29] Due to this ability, if they irradiated with a high intensity UV laser in the 254–360 nm range, they vaporize and form ions. From the early 90's several *cinnamic acid* derivatives were tested as sample matrices in different MALDI setups and substances like *proteins*, *polynucleotides*, *lipids*, etc. Nowadays, the MALDI technique uses many *cinnamic acid* derivatives[9], the most important ones are: *caffeic* **13**, *sinapinic* **14**, *ferulic* **15** and *α-cyano-4-hydroxycinnamic acids*[30] **16**, listed in *Figure 6*.

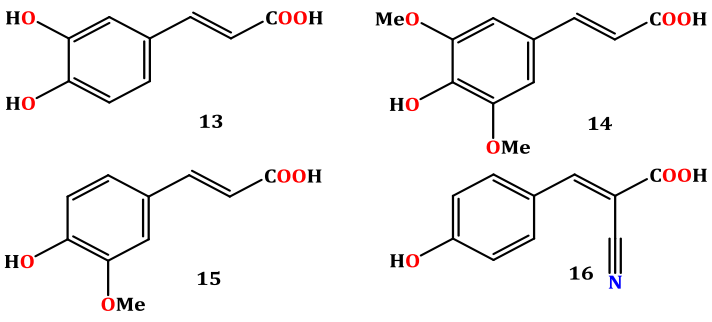


Figure 6 *Cinnamic acid* derivatives used as general sample matrices in MALDI

During the last 20 years *cinnamic acid* and its derivatives entered the spotlight again. Several pharmacological studies were carried out. Besides, the structure-forming properties were studied by some researchers as well.

The worldwide production of *cinnamic acid* is in the range 9–10 tons per annum, well indicating the importance of this compound.[4]

2.2 Natural occurrence and the synthesis of cinnamic acid and its derivatives

Cinnamic acid **1/E** and *cinnamates* are present in many plants, as has been mentioned above, as members of the *shikimate*[31] metabolism (it is a mayor pathway in the synthesis of aromatic *amino acids*) and the *phenylpropanoid* biosynthetic pathway.

2.3 Synthetic ways to E-cinnamic acids

As it has been mentioned previously, there are several different synthetic methods for producing *cinnamic acid* **1/E**; however, all of the listed reactions lead to the thermodynamically more stable *E* isomer. The most important synthetic methods for the preparation of *E-3-phenylpropenoic acid* **1/E** are communicated in the followings.

2.3.1 The Perkin condensation reaction

As it has been mentioned before, this reaction takes place between *aromatic aldehydes* **7** and *acetic anhydride* **8**, in the presence of the *potassium* or *sodium salt* of the corresponding organic acid (many other bases can catalyze the reaction, like *carbonates*, *phosphates*, *sulfides*, *organic amines* and even *sodium* or *potassium* metals).[32,33,34] A possible mechanism is shown in *Figure 7.*

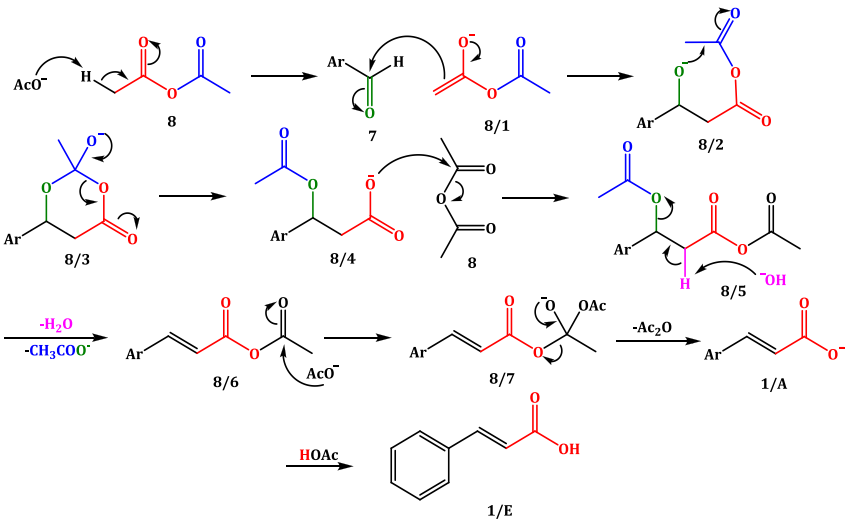


Figure 7 A possible mechanism of the original *Perkin* condensation[35]

The modified version of the *Perkin* reaction can be used for the synthesis of *α-substituted cinnamic acid* derivatives, where *acetic anhydride* **8** is the solvent and a

reactant at the same time, reacting with *acetic acid* derivatives (*phenylacetic acid* **17**) and usually amine bases is applied instead of inorganic salts.[36] The preparation of *E-2-phenylcinnamic acid* **18** (**E2Ph3Ph**) is shown in *Figure 8*.

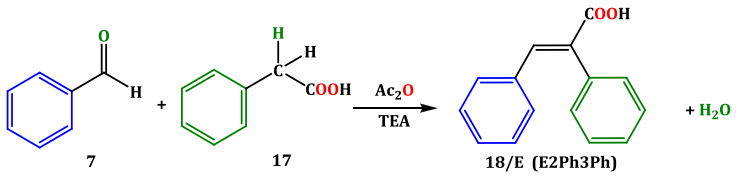


Figure 8 The modified *Perkin* condensation

The main disadvantage of the *Perkin* reaction is the formation of *resin*-like side products for *heterocyclic aldehydes*, and the low yield, when the *aromatic aldehyde* contains one or more electron donor groups.

2.3.2 The Claisen-Schmidt condensation reaction

The *Claisen-Schmidt* condensation[37,38] is a *cross-aldol* reaction taking place between several types of *aldehydes/ketones* and *carbonyl* compounds like *esters* in the presence of a strong base like *sodium*.

The formed *methyl cinnamate* can be converted to the *acid* with *ester* hydrolysis under alkaline conditions, followed by the addition of a strong acid (e.g. HCl). The overall yield is more than 90% for several *aromatic aldehydes*.

2.3.3 The Knövenagel-Döbner condensation reaction

One of the most effective synthetic methods, which can lead to *cinnamic acids* **1/E**, is the *Knövenagel* condensation[39] or "*malonester synthesis*" and its modified version, which is developed by *Döbner.*[40] The *Knövenagel* condensation works with *esters* (*Figure 9*)[41] e.g. *diethyl malonate* **19**, which loses an α *hydrogen* and forms a *carbanion* **19/1** in the presence of *piperidine* **20**, acting as the base, yielding *diethyl-benzylidene malonate* **21**. The most problematic part of this reaction is the heat treatment, at least 135–150°C is needed for the complete decarboxylation.[24,42]

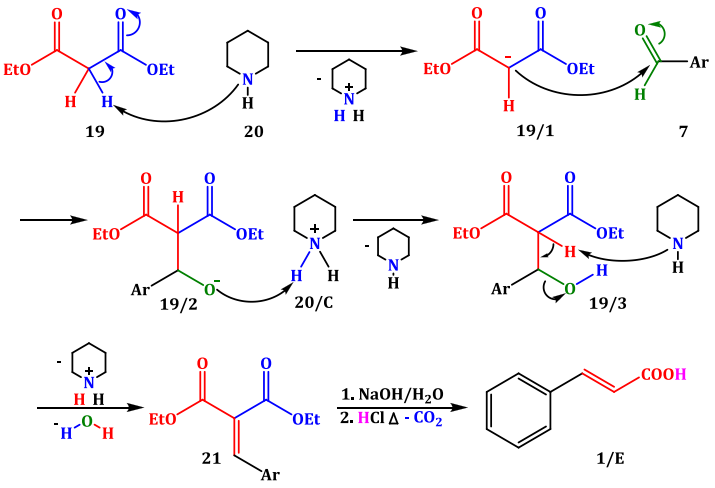


Figure 9 The mechanism of the *Knövenagel* condensation between *diethyl malonate* (**46**) and *benzaldehyde* (**7**)

With the *Döbner*[40] modification, *cinnamic acid* **1/E** can be synthesized directly from *malonic acid* **22**, dissolved in *pyridine* **43**, where the dehydration and the decarboxylation steps from **24**, can be carried out simultaneously in one pot, applying two temperatures during the synthesis. The catalyst in the *Döbner* variant is *piperidine* **20** as well; however, it works in the form of *piperidinium ion* **20/C**[43,44] (*Figure 10*).[45] The advantage of this reaction is the very high yield (in most cases 90–100%).

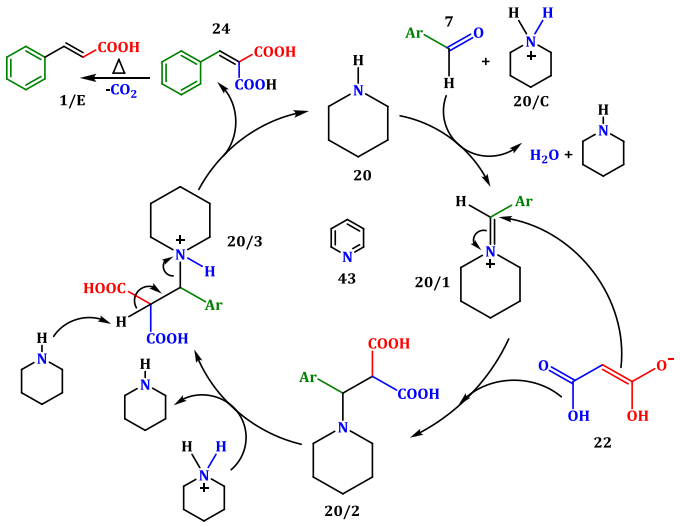


Figure 10 The catalytic cycle of *piperidine* (**47**) in the *Knövenagel-Döbner* condensation

2.3.4 The Heck coupling reaction

The so-called *Heck* coupling reaction (or *Heck-Mizoroki* reaction)[46,47] takes place between *alkyl* or *aryl halides* and *olefins* in the presence of *palladium-organo-phosphorous* complexes and a base. The *alkene* has to contain at least one *hydrogen* and

it performs better, if it contains EWG, e.g., esters like *methyl acrylate*, yielding *cinnamoyl esters*.

2.3.5 Microwave-assisted synthesis of cinnamic acid derivatives

The importance of microwave-assisted syntheses is growing very fast nowadays, due to the benefits of the method like higher yields, significantly shorter reaction time (1–30 min), and the solvent-free conditions in some cases, due to the "*superheating*" effect.[48] Most of the synthetic methods leading to *cinnamic acid* derivatives, mentioned earlier, can be carried out under microwave conditions (2.45 GHz), like the *Döbner* condensation in *water* with *phase transfer catalyst* or in the presence of PPE under solvent-free conditions.[49,50]

2.4 Synthetic ways to *Z*-cinnamic acids

As it has been shown on several examples above, the thermodynamically stable *E-cinnamic acid* derivatives can be easily achieved through various methods. With some modifications, one can shift the isomeric ratio towards the less stable *Z-cinnamic acid* isomer; however, the synthesis of the clean *Z* isomers is more complicated. A few general examples are listed in the followings.

2.4.1 Photoinduced isomerization of E-cinnamic acids

Molecules with conjugated electronic system (like **1/E**) can be excited with UV *photons*. The excited molecule becomes very reactive, but it can be relaxed in nonchemical ways like visible *photon* emission (*fluorescence* or *phosphorescence*) or trigger thermal reactions. However, the quantum efficiency is around only 0.2 in most of the organic reactions. On the other hand, in most cases, the excited molecule starts chemical reactions like rearrangement by isomerization, which is a useful tool to produce *Z-cinnamic acids* from the more easily available *E* isomers.

The isomerization of *E cinnamic acid* **1/E** (*Figure 11*) takes place without bond cleavage via a one-bond flip (OBF) around the double bond (diabatic photochemical reactions)[51], where the rotational barrier is in the 50–60 kcal/mol range; thus, the UV light (e.g., at 254 nm) has enough energy to induce the isomerization (112 kcal/mol).[52]

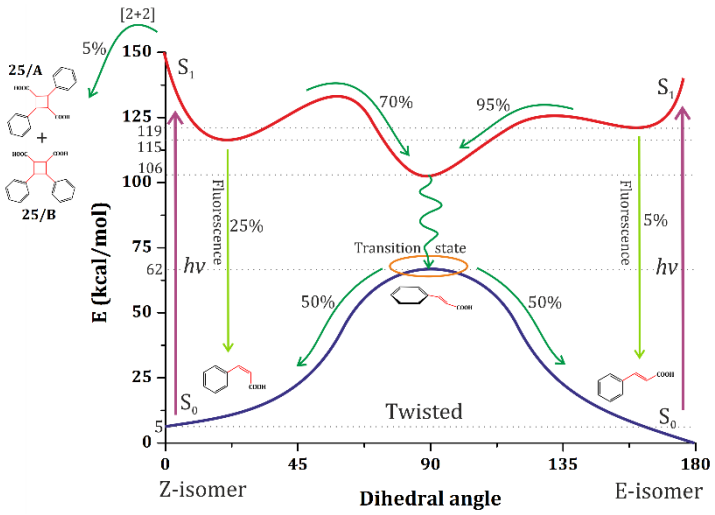


Figure 11 Estimated potential energy surface of the photoinduced processes with *E-Z cinnamic acids* in the UV range (based on *stilbene* **47**)

Both *E* and *Z* isomers have a wide absorption band in the UV range; thus, when they are irradiated with a *mercury* lamp (254 nm emission maximum) an $S_0 \longrightarrow S_1$ transition takes place, where the excited molecule can freely rotate around the double bond. There are two main possible ways to return to the ground state: the *fluorescent* way and the non-radiative relaxation. As one can see, the non-radiative way (green wavy arrow) is more frequent, and one gets even amounts of the isomers due to statistical consideration; however, the experimental data show a different result. The probability of the different types of relaxations is not equal; thus, one gets different isomer distribution: for *cinnamic acid* **1/E**, 75:25, *Z:E* ratio was found after 72 h irradiation at 254 nm.[52,53] On the other hand, at higher concentrations [2+2] topotactic photocycloaddition takes place, and the two possible dimers **25/A** and **25/B** slowly accumulate in the reaction mixture.[54] Under 0.1 M concentration, the formation of the dimers was negligible, and only the *E-Z* isomerization took place.

The choice of solvent is critical in photochemistry as well. There are two different types, the reactive (*water, alcohols, aromatics, alkyl halogenides*), absorbing UV light and the unreactive (*alkanes, esters*), which do not absorb in the UV range. For photoisomerization, the use of unreactive solvents is recommended.

2.4.2 Z-selective catalytic hydrogenation of alkynes

A very general method for the preparation of *Z-cinnamic acid* **1/Z** is the hydrogenation of *phenylpropiolic acid* **26** with *Lindlar's* catalyst (*Figure 12*), which is a

Pb-poisoned heterogeneous *palladium* catalyst[55] (usually 5% *Pd* by weight) deposited on a support of low surface area like $CaCO_3$ or $BaSO_4$. During the hydrogenation, the catalyst is further modified with *quinoline* **27** or with *3,6-dithia-1,8-octanediol* **28**[56], therefore the first equivalent of *hydrogen* enters via *syn* addition to produce *Z-alkenes*.

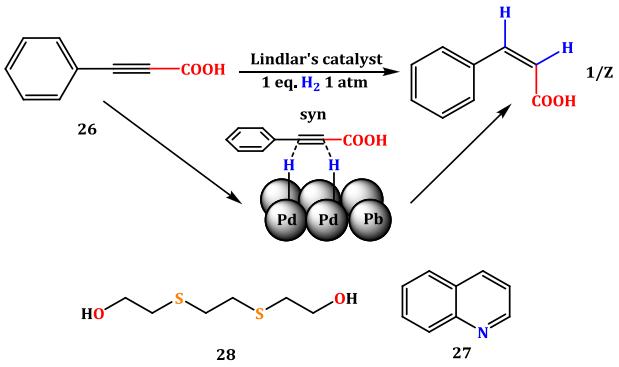


Figure 12 The mechanism of the *Lindlar*-type hydrogenation

The yield was quite good, more than 90% and the *Z:E* selectivity was around 95:5.[57]

An interesting variant of the *Lindlar* procedure is the so-called *P-2 nickel* catalyst, which is a Ni_2B powder, (usually generated *in situ*), which shows high *Z* selectivity for *alkyne* hydrogenation.[58]

2.4.3 The Horner-Wadsworth-Emmons (HWE) olefination

The HWE method[59,60,61], which is the offspring of the famous *Wittig* olefination[62,63], is one of the most favorable techniques for the stereoselective synthesis of *olefinic esters*. The mechanism is similar to that of the *Wittig* reaction, but the reagents are *phosphonates* instead of *phosphoranes*, and the *alkyl* group should contain an EWG, e.g., *ester* or *nitrile* groups. The different *phosphonates* can be synthesized utilizing the well-known *Michaelis-Arbuzov* rearrangement[64,65,66], where an *alkyl halide* reacts with organic *phosphites* to give the desired *phosphonate esters* like *trimethyl phosphonoacetate* **29**. In this reaction, the *carbanion* is generated with a strong base (NaH, *t*-BuOK, KHMDS **32**) under cryogenic circumstances (−78°C), usually in *ethers* like THF in the presence of complexone like *18-crown-6* **31**. The stabilized *carbanion* can react with *carbonyl* compounds like **7** via nucleophilic addition, to give *erythro* **29/4** and *threo* **29/3** *betaines*, being the rate-limiting step (*Figure 13*).[67] The *oxaphosphetanes* are formed as well here, and the elimination of **29/5** and **29/6** leads to *E* and *Z olefinic*

esters **30/E** and **30/Z**. The rate of the elimination will determine the isomer distribution, which is around 7:3 (*E:Z*) for *cinnamoyl esters* if the **R-** group is small (*methyl* or *ethyl*).

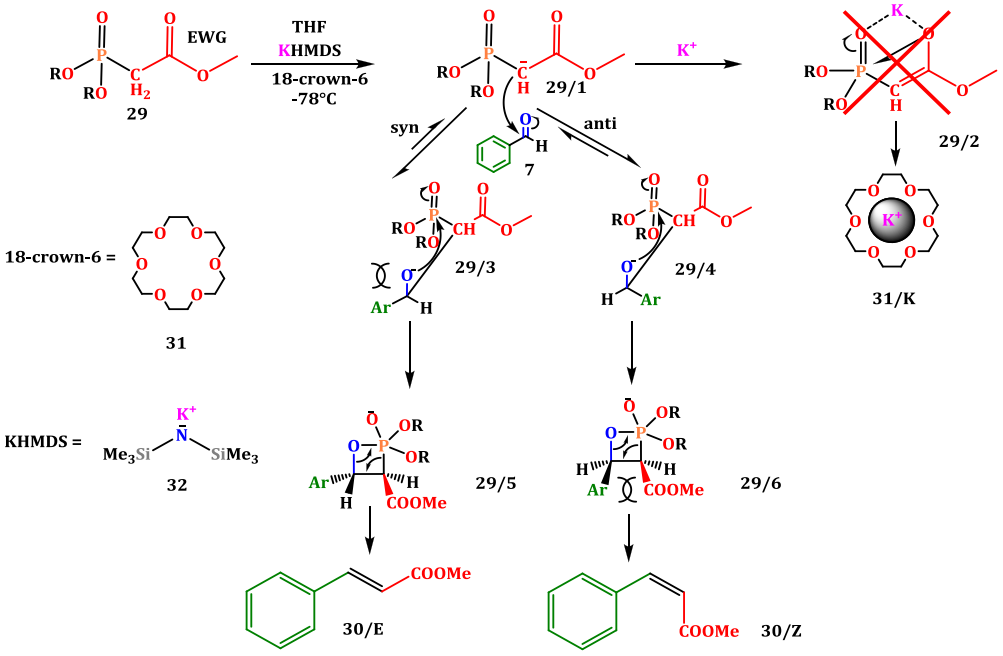


Figure 13 The working mechanism of the *Horner-Wadsworth-Emmons* condensation reaction

If the *phosphonate* has large side group, there will be a steric hindrance between the *phosphonate* and the ring of the *aldehyde*, making the *syn* addition sterically unfavorable and this will facilitate the *anti* route, which provides the *Z* isomer. Four basic *phosphonate* reagents can be used for the stereoselective preparation of *Z-cinnamic acid esters*: the *Still-Gennari*[68] **33** *bis-trifluoroethyl(R)phosphonoacetate*, the *Breuer*[69] ***34*** *dimethyl-dioxaphospholane-(R)acetate*, the *Patois-Savignac*[70] **35** *dimethyl-diazaphospholidine(R)acetate* and the *Ando*[71] **36** *bis-o-cresyl(R)phosphonoacetate* (*Figure 14*).

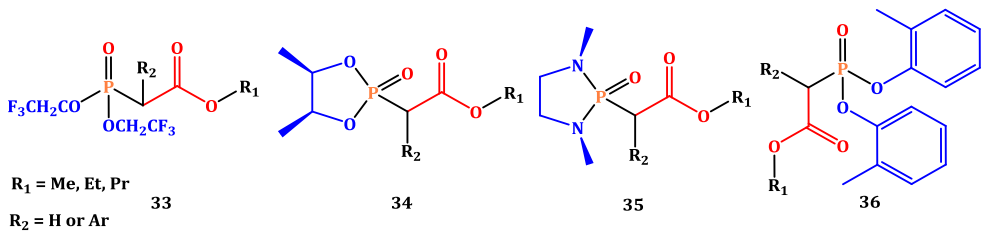


Figure 14 HWE-type *phosphonates* for *Z*-selective olefin synthesis

The HWE reaction can be applied for the synthesis of the *α-substituted cinnamic acids* **18/E-Z** as well, with the same reagents, if the **R2** group is an aromatic ring. From

the modifications listed above, the *Still-Gennari* **33** and the *Ando* **36** variants show the best *Z* selectivity for aromatic *aldehydes*.

2.4.4 Z-selective olefin metathesis

The *olefin metathesis* entails the redistribution of fragments of *alkenes* by the scission and regeneration of *carbon–carbon* double bonds. In spite of impressive advances in the development of catalytic *olefin metathesis* reactions[72], the lack of stereoselective access to *Z alkenes* represented a significant shortcoming in *olefin metathesis* for decades. Most of the metathesis reactions afford either mainly *E alkenes* or a mixture of the two isomers. Very recently, through the development of new *Mo-* or *W-alkylidene* complexes and *Ru-carbene* complexes, a general access to *Z alkenes* in high to excellent stereoselectivities through *olefin metathesis* was finally realized.[73] As one of the earliest demonstrations of the this new synthetic utility of selective *ring opening cross-metathesis* (ROCM) reactions, the *Crowe* group reported the first example of *Z*-selective *cross-metathesis* of *acrylonitrile* **38** and various terminal *olefins* like *styrene* **37**, catalyzed by *Schrock* catalyst **39** shown in *Figure 15.*[74]

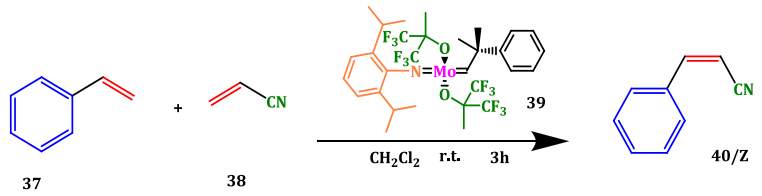


Figure 15 *Z*-selective cross-metathesis of *acrylonitrile* (**82**) with terminal *alkenes* like *styrene* (**81**)

The obtained *cinnamonitrile* **40/Z** can be hydrolyzed under alkaline or acidic conditions to access *cinnamic acids* (**1/Z**).

2.5 Synthetic ways to α-substituted cinnamic acids

The considerations concerning the synthesis of the *α-substituted cinnamic acids* are very similar to those of *cinnamic acid* (**1/E**); however, the isomer distribution can be adjusted with the modification of the reaction time, temperature and with the reagent type as well.

2.5.1 The modified Perkin and other relevant condensations

The first method for the synthesis of *α-phenylcinnamic acid* was carried out with the modified *Perkin-Oglialoro* condensation between *benzaldehyde* derivative **7/R** and

phenylacetic acid derivative **17/R** (or the corresponding *alkali salt* **17/R-Na**), catalyzed with TEA in Ac_2O[75] (*Figure 16*). The *aromatic aldehyde* and the *aromatic acetic acid* derivative can contain a *heterocycle* as well (*furyl-, thienyl-*)[76]; thus, a wide variety of *α-substituted cinnamic acids* can be prepared with these methods, obtaining both isomers with high yields.

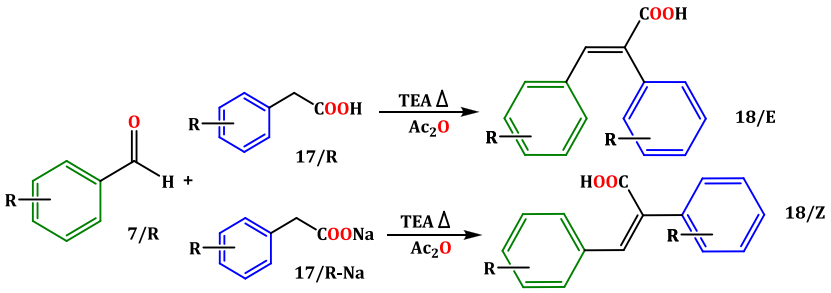


Figure 16 The modified *Perkin* and *Oglialoro* condensations

Other condensation reactions such as the HWE method (*2.4.3*) can be applied as well to produce *α-substituted cinnamic acids*, if we are able to create the corresponding *phosphonates* **41** (*Figure 17*) for the synthesis, where the R_2 group can be *benzyl-* **42**, *pyridyl-* **43**, *furyl-* **44**, *thienyl-* **45** or *selenophenyl-* **46**.

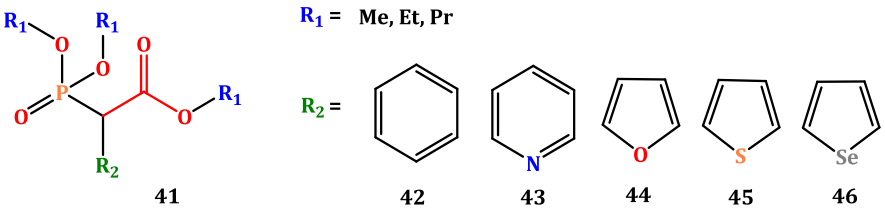


Figure 17 HWE *phosphonate* for the synthesis of *α-substituted cinnamic acid* isomers

2.5.2 Synthesis of α-substituted cinnamic acids via addition reactions

Various other methods for the preparation of *α-phenylcinnamic acid are already mentioned*, however each described methods are based on condensation. Other relevant reactions can be found in literature, like the radical *Kharasch*[77] addition of *halomethanes* **48** to *stilbene* **47/Z** trough **49** (*Figure 18*)[78] or the *hydrocarboxylation* of *diphenylacetylen* in the presence Et_2Zn and CsF[79], can be found; however, the produced molecules will carry the geometry of the starting material; thus, the selectivity is predetermined in these reactions and the preparation of the starting compound is difficult is some cases.

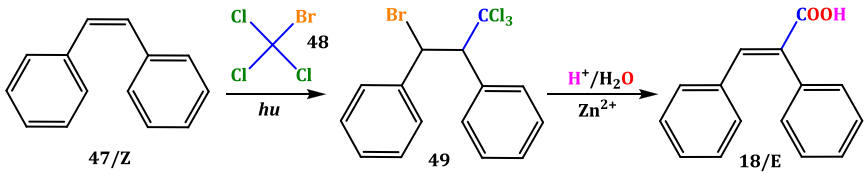


Figure 18 The radical addition of *bromotrichloromethane* (**48**) to *stilbene* (**47/Z**), than the hydrolysis of (**49**) under acidic conditions

2.6 Secondary chemical bonds and the self-assembling capabilities of cinnamic acids

2.6.1 About secondary chemical bonds

There are several well-defined secondary chemical bonds that can hold small and large molecules together. Each secondary chemical bond is based on the electrostatic or dipole-dipole interactions between strong or permanent dipoles of neutral molecules and ions respectively. They can be categorized by the type and the strength of the appropriate interaction as follows:

- *hydrogen bonds* (inter- and intramolecular)[80]
  - First-order *hydrogen bonds* (1.9–38.6 kcal/mol)[81]
    - F–H…F (38.6 kcal/mol)
    - O–H…N (6.9 kcal/mol)
    - O–H…O (5.0 kcal/mol)
    - N–H…N (3.1 kcal/mol)
    - N–H…O (1.9 kcal/mol)
  - Second-order *hydrogen bonds* (1–6 kcal/mol)[82,83]
    - O–H…X (X = S, Se,Te,Hlg)
    - N–H…X
    - C–H…F
    - C–H…O
    - C–H…N
    - C–H…S
  - *Dihydrogen bond* (between *hydrides* and positively polarized *hydrogens*, 1–25 kcal/mol)[84,85]
    - *alkali*–H…H–X (X = O,N,S,C,Hlg)
    - B–H… H–X

        - Al–H… H–X
        - Ge–H… H–X
        - *transition metal*–H… H–X
- *van der Waals* forces[86]
    - *Keesom* (permanent *dipole–dipole* interaction)
    - *Debye* (*induced–dipole* interaction)
    - *London* dispersion force (between nonpolar molecules)
- Other weak interactions[87]
    - *π–π stacking* (between *aromatic rings*, 0.5-3 kcal/mol)[88]
    - *C–H…π stacking*
    - *cation*–π bonding
    - *anion*–π bonding
    - *halogen bonding* (similar to *hydrogen bond*, A…X–A)[89]

The interactions listed above are very important in defining the physical properties of materials, like the melting and boiling point (*water*[90] vs. *hydrocarbons*) or even the viscosity (*sugars*) and solubility, the structure forming properties (multiple *hydrogen bonds* in DNA[91] and *polymers*), the reactivity (intramolecular *hydrogen bonds* in *enols*), etc.

2.6.2 Possible interactions between cinnamic acids

Many of the mentioned secondary bonds are possible among *cinnamic acid* derivatives; therefore, they are good models for observing various aggregations types in all three thermodynamic states, or even in the adsorbed phase.

A. First-order *hydrogen bonds* between *cinnamic acids*

One of the strongest *hydrogen bond* is the O–H…O type, which is very general between *carboxylic acids*. If the formation of the bond is not hindered by large substituents and the *carboxylic* group is in the *syn periplanar* conformation, a so-called cyclic or double *hydrogen bond* can be established. The structure of the cyclic dimer of two *cinnamic acid* monomers **1/E** is displayed in *Figure 19*, where the *hydrogen bond* is illustrated as dashed green line.

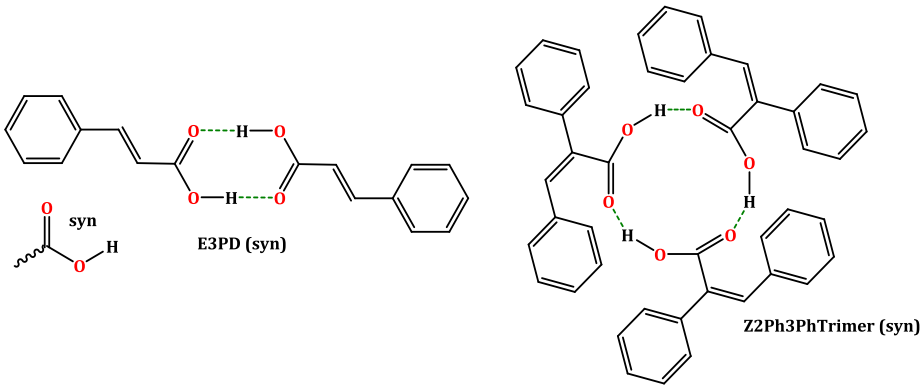


Figure 19 Cyclic *double* and triple O–H...O *hydrogen bonds* between two *cinnamic acid* (**1/E**, **E3PhD**) and *Z-α-phenylcinnamic acid* (**Z2Ph3PhTrimer**) molecules

The double *hydrogen bond* is significantly stronger than two separated O–H...O interactions, the binding energy is usually in the 14–17 kcal/mol range[92,93] instead of 10 kcal/mol that can be expected from the energy of a single O–H...O bond (5–6 kcal/mol). The short bond length indicates the strength of the interaction, it is in the 1.65–1.81 Å[15] range for various *cinnamic acid* derivatives[94] that is very close to the average distances, which can be measured among covalently bonded atoms. Due to the strength of these bonds, almost every *cinnamic acid* derivative forms dimers in the solid state uniformly[95], and the cyclic dimers will stay together either in aprotic solvent[82] or sometimes in the gas phase[96] as well. For *α-substituted cinnamic acids*, the formation of *hydrogen-bonded* aggregates via the *carboxylic* groups, did not stop at the dimer, but *hydrogen-bonded* cyclic trimer was also formed in the solid state and even in concentrated solution.[97]

B. Second-order *hydrogen bonds* between *cinnamic acids*

As it has been mentioned at the secondary chemical bonds, many weaker *hydrogen bonds* are possible between *aromatic protons* and *heteroatoms*. We have already seen the structure of the **E3Ph**(*syn*) dimer, but the *anti* conformer of the *carboxylic* group also exists[98], and with the *anti* conformer, the formation of the *cyclic dimer* is prohibited, due to inappropriate bond lengths and angles. In this case, the aromatic C–H...O interaction became favorable in the solid state[95] (*Figure 20*); however, the incidence of the *anti* conformer is rare (1–5%), due to the higher energy content (1–2 kcal/mol over the *syn* conformer).[99] The energy of the (aromatic)C–H...O *hydrogen bond* is in the 0.5–1.2 kcal/mol range, and the bond length varies from 2.4 to 2.6 Å.[100] Two different structures can be seen in *Figure 20* for the C–H...O interaction (represented by orange dashed lines) between the *anti* conformers of *cinnamic acid* (**1/E**, **E3Ph**).

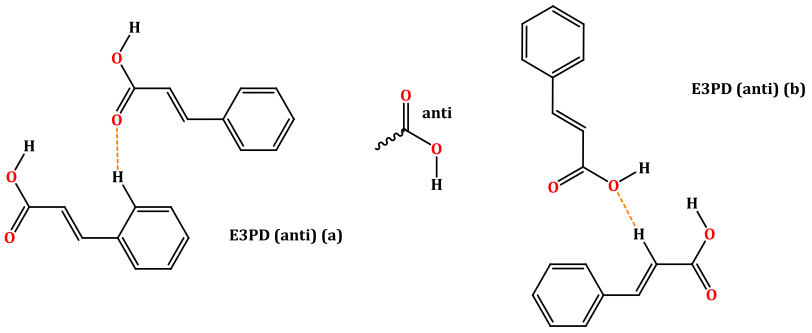


Figure 20 Various C–H…O interactions between *cinnamic acid anti* conformers **E3PhD**-*anti* (**a**) and (**b**)

A little bit stronger *hydrogen bonds* may be formed between *heteroaromatic* rings with cyclic C–H…O and C–H…N interactions[101], and data are found for the weaker C–H…S bond as well. Some other types of *hydrogen bonds* should be mentioned here, for instance, the aromatic C-H…F bond between *fluoro substituted cinnamic acids.*[82] As it has been mentioned before, these interactions are weaker compared to the "traditional" *hydrogen bonds*, due to the slightly polarized *aromatic protons*[10]; however, the cyclic *hydrogen bond* among *heteroaromatic* rings or substituted *phenyl* rings, can be as strong as 1–6 kcal/mol, which is not negligible in the solid state, or even in the liquid phase.[102] Due to the lower bonding energies, the average bond length is quite long: 2.4–3.5 Å.[103,104] A possible cyclic C–H…O agglomerate, consisting of *E-3-(2-furyl)propenoic acid* dimers **E32FD**, and a *fluorophenyl* substituted **Z24FPh34FPhD** are shown in *Figure 21* (where the weak *hydrogen bonds* represented as pink dashed lines).

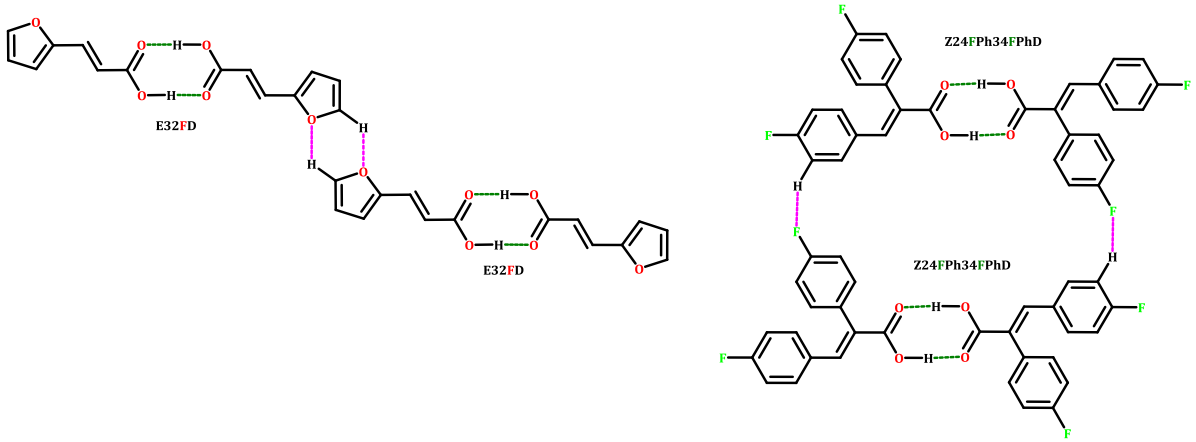


Figure 21 Example for the *cyclic heteroaromatic* C–H…O *hydrogen* bond and substituted aromatic C-H…F *hydrogen bond*, formed between dimers

C. *π-stacking* interactions between *cinnamic acids*

The weakest interactions observed for *cinnamic acids* are the *π–π stacking* and the C–H…π close contact. It is known that these kinds of interactions play role in the

conformation of *proteins* and other biologically important molecules[105], and they determine the crystal structure of aromatic compounds. There are several different theories about the nature of the *π-stacking* interaction, but the most convincing is the electrostatic model[106], which found direct correlation between the structure of different *π-stacked* compounds and the substituents of the aromatic ring.[107] The three basic models of the *π-stacked* structures are displayed in *Figure 22*.

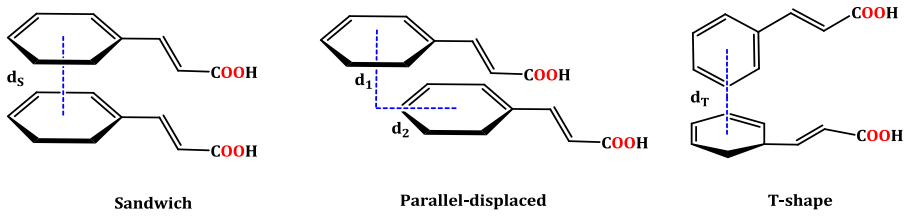


Figure 22 The three basic *π-stacking* interactions presented on *cinnamic acid* (**1/E**)

The average bonding energy is between 0.5–3.0 kcal/mol and the order in ascending order for *benzene* derivatives is the following: *parallel-displaced*, 2.6 kcal/mol (d_1= 3.6 Å, d_2= 1.6 Å); *T-shaped* or (aromatic)C–H…π, 2.5 kcal/mol (d_T= 5.0 Å); and the *sandwich-type* interaction, 1.7 kcal/mol (d_S= 3.9 Å).[108] The *parallel-displaced* and the *sandwich-type* interactions are present in *cinnamic acid* agglomerations in the solid state, as can be deduced from crystallographic data.[95,109]

2.6.3 Self-assembling of various molecules over transition metal surfaces

As it was seen, several different types of interactions and bonds can be formed among *cinnamic acids*, even if they contain EWG substituent(s) on the aromatic ring, or when the aromatic ring is a *heterocycle*. With the use of metal surfaces, we can add another dimension to the organizing forces via the attraction between organic functional groups (mostly with groups containing O, N or S atoms) and transition metals (e.g., *Pt*, *Pd*, *Au*, *Ag*)[110]. Many examples can be found in the scientific literature describing the chemisorption of organic compounds over metal surfaces via their heteroatoms. There are two types of chemisorption, which may involve self-assembly:

- The stronger covalent adsorption taking place via the cleavage of *hydrogen* from the terminal –OH, –COOH, –NH_2 or –SH functional groups.[111] The chemisorbed layer is usually called *self-assembled monolayer* (SAM). The formation mechanism of the covalent-type SAM is seen in *Figure 23*. The first step (**I.**) is the physisorption of the molecules (presented as grey sticks where the dark "*head*" can be –COOH, –OH, –NH_2 or

–SH functional groups). The second step (**II.**) is the formation of the covalent M–O, N, S bond via cleave the *hydrogen* off the functional *head-group* (drawn as red "*heads*"). In a further step (**III.**), the chemisorbed molecules start to assemble from parallel to perpendicular fashion, organized by second-order chemical bonds among the side groups and chains. The final state (**IV.**) is what we call SAM, a well-ordered carpet-like structure, covering the surface of the metal and completely changing its electronic and chemical properties.

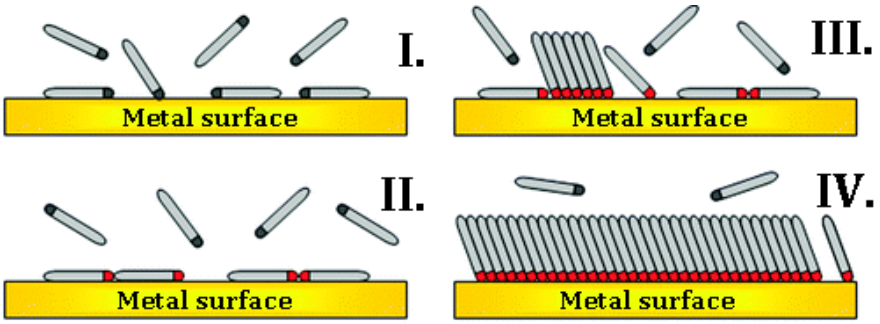


Figure 23 The common mechanism of covalent *self-assembling* over metal surface[112]

- The weaker coordinative adsorption via R–O–R, R=N–R, R–S–R groups without bond-cleavage[111], where the *heteroatom* can be located in an aromatic ring or surrounded by *alkyl* neighbors. The formation of the SAM follows the same route as shown in *Figure 23*, and oriented nearly perpendicular to the surface for simple *heteroaromatics* like *pyridine* **43**[113,114] and *thiophene* **45**.[115] If the substance contains more than one coordinating *heteroatoms*, relatively far from each other, the formation of the SAM may stop at stage **II.**, and the orientation stays parallel to the surface.[115,116]

The SAM layers are widely used for surface modification mainly for anticorrosion coating[117], changing the wetting and adhesion properties, achieving chemo- and biocompatibility[118], or for the creation of nanostructures made for electronic use, like sensors or conducting polymers (usually made by polymerization of the adsorbed molecules).[119,120]

If the SAM has functionalized tail group, another layer of molecules can be attached to it and the procedure can be repeated again to grow a multilayer. This technique was developed by *DuPont*[121] to build uniform coating on *glass*, and is known as *Layer-by-Layer* (LbL) method. The original procedure uses electrostatic forces to build the layers by altering the positively and negatively charged particles.[122] However, many other

interactions can hold layers together like *hydrogen bonds* or dispersion forces.[123] The main advantage of the procedure is its simplicity. The layers form spontaneously, i.e., simple dipping (*dip-coating*) and washing steps are only required to form the multilayer.

As was seen above, heteroaromatic molecules can form SAMs, especially the *sulfur*-containing derivatives like *thiophene* **45**; therefore, *heterocyclic cinnamic acids* have great potential in forming SAMs over *gold* or *silver* surfaces, and through intermolecular *hydrogen bonds* mentioned earlier, the formation of LbL films are possible as well.

3. THE MAIN AIMS OF THE DISSERTATION

The goals of the work leading to this dissertation are as follows:

- finding a cheap, successful synthetic way to prepare stereochemically pure *E-2 and/or 3-(furyl/thienyl/pyridyl/selenophenyl)cinnamic acids*, because some of them have very high price on the list of chemical suppliers;
- elaborating an expedient and affordable way for the preparation of hardly accessible *Z-3-(furyl/thienyl/pyridyl/selenophenyl)cinnamic acids*, because none of them are available commercially;
- studying the *self-assembling* features of the above molecules and many of their substituted varieties in solution, the solid state and the adsorbed phase on metal surfaces (*gold* and *silver*), identifying the possible secondary interactions, with the combination of various spectroscopic measurements (^{1}H NMR, IR) and surface inspection techniques (SEM, AFM, IRM);
- complementing the experimental results with computational methods at high theoretical level (geometry and energy calculations for the aggregates, including the computation of precise energies of the secondary bonds and the full conformational analysis of the individual molecules and occasionally the dimers as well).

4. EXPERIMENTAL PART

4.1 Materials

4.1.1 The molecules studied

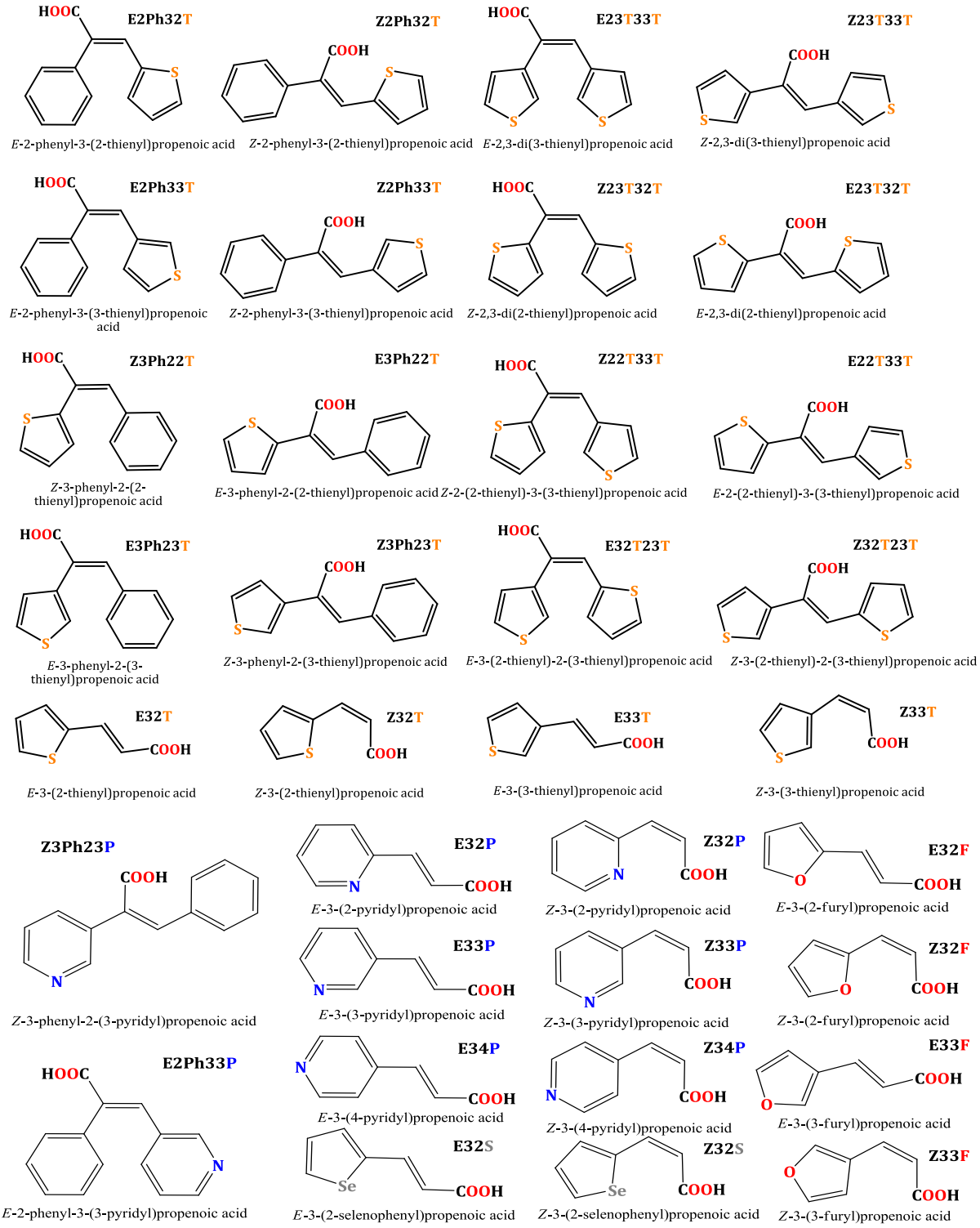


Figure 24 The structure of the studied *cinnamic acid* derivatives, substituted with one or two *heteroaromatic* ring(s)

The structure of the studied materials of this dissertation can be seen in *Figure 24.*

4.1.2 Solvents for NMR spectroscopy

- **DMSO-d_6 (*hexadeutero dimethyl sulfoxide*)** (99.96 atom% D, 0.03% TMS) (Sigma-Aldrich, NMR solvent)
- **D_2O (*deuterium oxide, heavy water*)** (99.9 atom% D, 0.01% DSS) (Sigma-Aldrich, NMR solvent)
- **$CDCl_3$ (*deuterochloroform*)** (99.8 atom% D, 1 V/V% TMS) (Sigma-Aldrich, NMR solvent)
- **CD_3OD (*tetradeuteromethanol*)** (99.8 atom% D, 0.03% TMS) (Sigma-Aldrich, NMR solvent)

Other reagents and supporting materials applied for the experiments are listed in the *Appendix.*

4.1.3 Polycrystalline metal surfaces

- **Gold and silver surfaces**

The polycrystalline metal surfaces were created by the *Pulsed Laser Deposition* (PLD)[124] technique. This method uses incident focused UV-laser beam to vaporize (ablate) bulk metal, which is placed in a vacuum chamber, and the liberated plasma particles condense on the substrate, placed in front of the plume (*Figure 25*). With the control of the pulse length, the frequency and the energy of the laser beam, very thin metal layers can be created on several substrates (usually *glass*, *quartz* or various *metals*). The common PLD setup is displayed in *Figure 25*. For the measurements 30×30mm *glass* microscope slides were applied as substrates, and 99.99% *gold* and *silver* (Good Fellow) plates as targets. The applied laser was a *krypton fluoride* (KrF) excimer laser (λ = 248 nm), focused on the metal target in a 5 mm^2 spot. The pulse length was 18 ns, the repeating frequency was 10 Hz and the energy of a single impulse was 106 mJ/pulse.

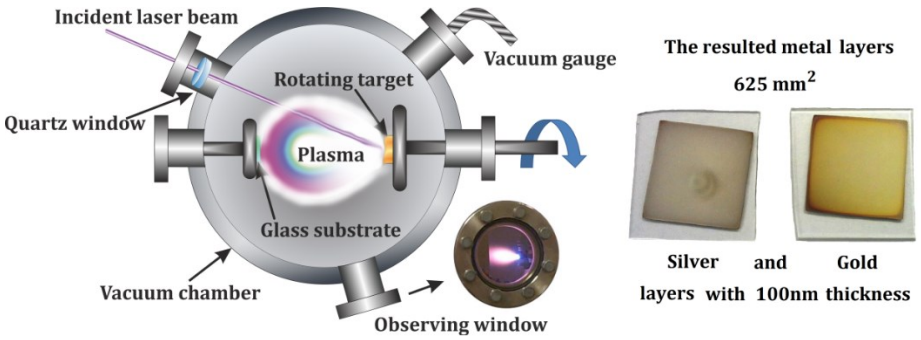


Figure 25 Schematic representation and the working mechanism of a common PLD setup and the resulting metal layers

Each metal layer was created with 15,000 pulses, and the resulting layer thickness was around 100 nm. The area of the created metal surface was 25×25 mm (625 mm^2) in every case. The homogeneity of the metal layers was checked by *scanning electron* (SEM) and *atomic force microscopies* (AFM). At the atomic level, it was found to be acceptable for topological studies (1–2 nm peaks between the huge 1–3 μm droplets, coming from the perimeter of the ablation craters, where the heat is no longer enough to vaporize the metal).

- **Creating the self-assembled layers**

Each metal layer was treated with boiling DMF to remove the oxide cover from the surface[125], and then, they were washed in a *chloroform* bath, combined with ultrasonic cleaning (50 W US, 20 kHz, 10 s) to remove all impurities before the experiments. The slides with the deposited metal layers were dipped (*dip-coating* method) into the 0.1 M solution of the proper *cinnamic acid* derivative dissolved in *chloroform* (0.1 M is far from the saturation point of the five membered derivatives, but the *pyridyl*-derivatives are barely soluble in chloroform, so the selected concentration for them is 0.001 M) for 48 h to create the *self-assembled layers*. After the treatment, the layers were washed with 3×5 ml of *chloroform* to prevent the formation of non-bonded crystals over the *self-assembled layers.*

4.2 Instrumental methods

4.2.1 Analytical and structure analysis methods

- **Gas chromatography (GC)**

All gas chromatographic measurements were carried out on a HP-4890 *gas chromatograph* equipped with a *flame ionization detector* (FID). A HP-5 type *polysiloxane*

cross-linked (Me:Ph, 95:5) 50 m×0.32 mm capillary column with 25 µm internal diameter and 0.52 µm film thickness was applied. The instrument was used in the splitless mode, and the column head pressure was 2 bars. The FID, was fed with 99.998% H_2 (Messer) and 99.995% compressed air (Messer). The carrier gas was 99.995% N_2 (Messer). Each measurement was carried out with 0.1 µl sample volume, injected with a 5 µl syringe. The results from GC measurements were calculated by the integration of the chromatogram.

- **Ultraviolet-Visible spectroscopy (UV-Vis)**

The electron excitation measurements were performed on a Shimadzu 1650-PC UV-Vis two-way spectrophotometer. Each measurement was recorded in a 10 mm *quartz* cuvette (Hellma) between 200–400 nm, with 1 nm resolution and 4 nm/sec scan speed. The reference was HPLC grade EtOAc in every case.

- **Infrared spectroscopy (IR)**

The *infrared* (IR) measurements were carried out on a Bio-Rad FTS 60-A FT–IR spectrometer. Each spectrum was collected between 4000–400 cm^{-1} with 4 cm^{-1} resolution and 256 accumulations. For the *α- thienyl substituted cinnamic acids* and for the **E32P**, **E33P**, **E34P** and **E2Ph33P** molecules, the spectra collected in solution as well using NaCl cuvettes with different light path lengths (the solvent was DMSO or *chloroform*, and the absorption bands of the solvent are removed from the spectra DMSO: 2900–3000 $^{-1}$cm, 1460–1380 cm^{-1}, 1100–925 cm^{-1}, chloroform: due to the very low absorptions of it, there is no need to remove the absorption bands). The solid samples measured utilizing the *attenuated total reflection* (ATR) mode, using a Harrick-Meridian split-PEA *diamond* ATR head and a liquid *nitrogen* cooled MCT detector with 20 kHz scanning speed. Each spectrum was baseline-corrected, and due to the vibrations of the *diamond C–C* bonds, the 2200–2000 cm^{-1} region was removed from every spectrum. On using other molecules, the "classic" pastille method was applied, where the sample is mixed and encapsulated with KBr using 1 wt.% of the sample molecule and 256 scan were collected using a DTGS detector.

- **Nuclear magnetic resonance (NMR)**

Nuclear magnetic resonance (NMR) measurements were performed on a Bruker Avance 500, 500 MHz spectrometer. For the ^{1}H and ^{31}P spectra, 128 scans were collected. The ^{1}H–^{1}H *Correlation Spectra* (COSY) measured with 1024×512 resolution and 16 to 64 scans were accumulated. The *nuclear Overhauser effect spectroscopic*

(NOESY) and *heteronuclear single quantum coherence* (HSQC) results were gained with the same parameters as was described for COSY. Each measurement was performed in deuterated solvents (described in the materials section) in Wilmad, 5 mm, 500 MHz NMR tubes at atmospheric pressure and 298 K temperature. Each spectrum was baseline- and phase-corrected. The t1, t2 noise and the contaminants, like solvent peaks or impurities are removed from the 2D NMR spectra to simplify the images.

- **X-ray diffractometry (XRD)**

Powder *X-ray diffraction* (XRD) patterns of the solid samples were registered in the 2Θ = 3–35° region on Rigaku Miniflex II and DRON-2 instruments, using CuKα (λ = 1.5418 Å) radiations in *Bragg-Brentano* geometry. Reflection positions were determined via fitting a *Gaussian* function. They were found to be reproducible within 0.05° (2Θ).

- **Melting Point (mp)**

The melting points of the solid materials (if the melting points cannot be found in the literature) are determined on an SRS Optimelt MTA100, automatic melting point measuring system, which uses the capillary *glass* tube method with optical analysis. Each measured or cited melting point data are in °C.

4.2.2 Surface inspection techniques

- **Infrared Microspectroscopy (IRM):**

The IRM technique was applied for the characterization of the adsorbed molecules on *gold* and *silver* surfaces (created by the PLD method). The measurements were carried out on a Bruker Vertex 70 IRM instrument, collecting the spectra in total reflection mode from the focused area of the proper surface (1–0.1 mm^2). A liquid *nitrogen* cooled MCT detector was used to collect observable intensities. Each spectrum was created from 1024 accumulated scans with 2 cm^{-1} resolution in the 4000–400 cm^{-1} region.

- **Scanning Electron Microscopy (SEM)**

The metal layers created via the PLD technique were inspected with *scanning electron microscopy*, to determine the homogeneity of the surface at the atomic level. The instrument used for the characterization was a Hitachi S-4700 device with varying accelerating voltages. The samples were fixed and grounded with a copper tape.

- **Atomic Force Microscopy (AFM)**

The topology of the adsorbed molecular layers on *gold* and *silver* surfaces was studied by *atomic force microscopy* using an NT-MDT Solver AFM microscope working in the "*tapping*" mode (the frequency of the cantilever was 278 kHz). The AFM needle (manufactured by Nanosensors, Inc.) was an SSS-NCH-type 15 µm long *silicon* needle with 10° half cone angle and 2 nm radius of curvature.

4.2.3 Reactors and preparative instruments

- **Microwave reactor system (MW)**

For the *microwave-assisted* reactions, a CEM Focused Microwave System (Model Discover SP) was applied. The system runs at 300 W, and the temperature was kept under max. 90°C for varying times, using a 100 ml rbf. equipped with an air-cooled condenser (open vessel method).

- **UV-irradiated reactor (UV)**

For the UV induced isomerization experiments, a semi-continuous UV-irradiated flow reactor was designed. The isomerization reactor was a standard 10 mm (4 ml), four-sided *quartz* cuvette with a Teflon plug (Hellma). The cuvette was attached to a 12 V DC membrane pump **M1** (ASF Thomas D-8217) and two 100 ml *glass* buffer vessels via Teflon tubing (3 mm external and 2 mm internal diameter). The flow of the treated solution was controlled by an electronic circuit based on an ATTINY 2313-P microcontroller unit (MCU) fed by stabilized 5 V power supply and the EEPROM programmed with an AVR−ISP MK II USB-programmer interface. All components of the electronic circuit are class I. electronic products (1% metal layer resistors: **R1-R3-R4**, high quality ceramic capacitor: **C1** and class I. trimmer potentiometer: **R2**). The MCU drives an IRFZ 24N MOSFET **Q1,** which switches on and off the pump at the proper time (*Figure 26*). The UV-source was a *mercury-vapor* lamp (Osram, T5 germicide light-tube, 135 mm, 4 W, G5) with an emission maximum at 254 nm. The maximum power at the emission line was 1 W. Each isomerization reaction was carried out in HPLC grade EtOAc (99.9%). The concentration of the treated solution was 0.05 M in all cases to avoid the formation of unwanted side products.

The schematics of the flow reactor and the electronic circuit are shown in *Figure 26* and *Figure 27*.

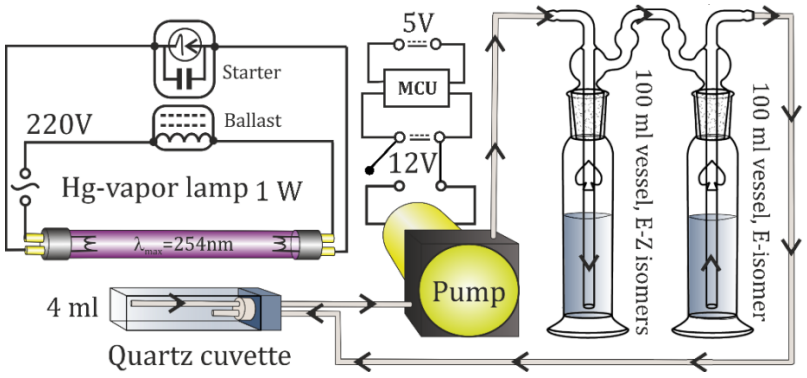


Figure 26 The schematic of the automated semi-continuous UV flow reactor

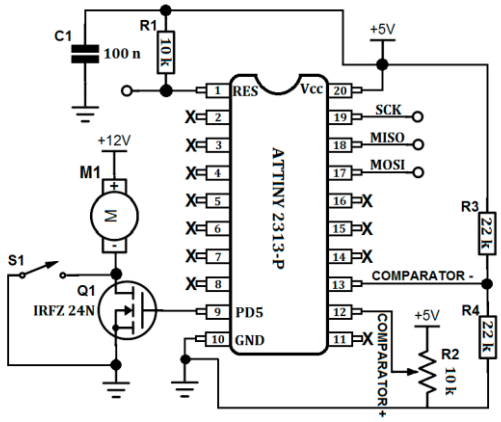


Figure 27 The circuit scheme of the flow controller unit

The C source code of the controller, driving the MCU, can be found in the *Appendix*.

4.3 General methods for the stereoselective synthesis of *E*-cinnamic acids

The short description of the tested methods will communicated in the following.

4.3.1 Preparation of 2-selenophenylcarboxaldehyde with Vilsmeier-Haack formylation

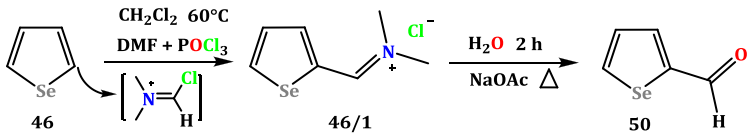


Figure 28 Formylation of *selenophene* (**46**) with the *Vilsmeier-Haack* method

The preparation of *2-selenophenecarboxaldehyde* **50** was carried out with optimized *Vilsmeyer-Haack formylation*. 1 eq. **46** dissolved in *dichloromethane*, then cooled, and 1 eq. of DMF and $POCl_3$ were added. The mixture was heated to 65°C in a sealed bottle for 1 hour; then, refluxed with aqueous *sodium acetate* for additional 2

hours. The reaction mixture was extracted with *dichloromethane* dried over sicc. $MgSO_4$, and then evaporated. The residue was vacuum-distilled at 100 mbar, and the fraction between 90–100°C was collected as colorless oil, which was **50**, 65% of the theoretical yield (99%, GC). The preparation of the *3-selenophenecarboxaldehyde* was skipped due to the too many reaction steps needed requiring high amount of *selenophene* **46**, which is a highly toxic and expensive compound.

4.3.2 Preparation of E-cinnamic acid with the Perkin condensation

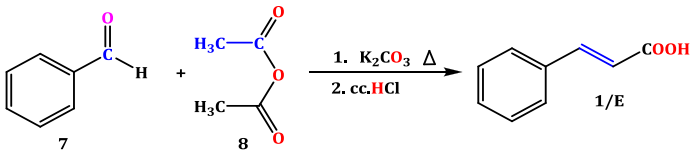


Figure 29 Preparation of *E-cinnamic acid* with *Perkin* condensation

The reaction was carried out with the condensation of *acetic anhydride* **8** and *benzaldehyde* **7** in the presence of K_2CO_3 with gentle heating for 2 h, following the literature method.[124] The mixture diluted with KOH solution, extracted with *diethyl ether* and the *aqueous* phase was acidified. The precipitate was filtered off and dried. The isolated yield was 81–92%. The solid material was recrystallized from MeOH:H_2O (1:3) applying activated *charcoal*. The product was pure **1/E** (99%, ^{1}H NMR).[126]

4.3.3 Preparation of E-ethylcinnamate with the Claisen-Schmidt condensation

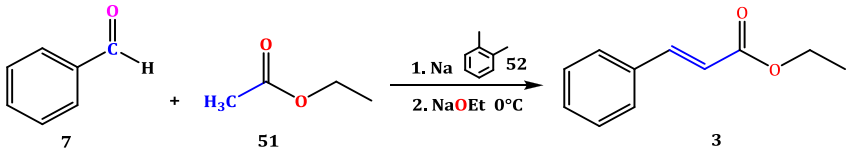


Figure 30 Preparation of *E-cinnamic acid* with *Claisen-Schmidt* condensation

The *Claisen* condensation was performed with a general method.[127] Clear *sodium*, melted in *o-xylene* **52** and parsed into small particles; then, the solvent changed to EtOAc **51**, containing catalytic EtOH. The mixture cooled to 0°C and *benzaldehyde* **7** was added dropwise. After all *sodium* reacted, EtOAc distilled off and the residue was vacuum-distilled under 100 mbar, the fraction between 180-190°C was collected as a colorless liquid, which was pure *E-ethylcinnamate* **3**. Yield: 60-65% (98%, GC). The ^{1}H NMR spectrum correlates well with the literature data.[128]

4.3.4 *Preparation of diethyl-2-benzylidene malonate with the Knövenagel condensation*

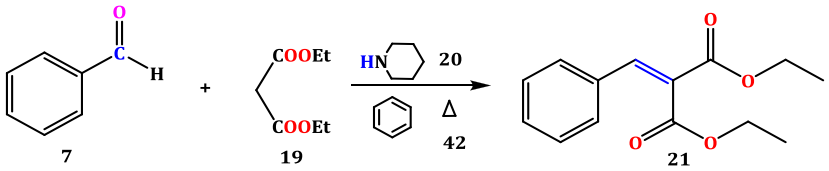


Figure 31 Preparation of *diethyl-2-benzylidene malonate* with the *Knövenagel* condensation

The preparation was carried out on a *Dean-Stark* water trap.[129] *Diethyl malonate* **19**, commercial *benzaldehyde* **7** (commercial *benzaldehyde* contains a few percent of *benzoic acid*, which is necessary for the reaction to work), catalytic amount of *piperidine* **20** and *benzene* **42** were refluxed for 10 h, until no more *water* was formed. The volatiles were removed with distillation and the residue was vacuum-distilled at 100 mbar, and the fraction between 220–230°C was collected as a light yellow oil (**21**).[130] Yield: 85–87% (98%, GC). The ^{1}H NMR verified the structure of **21**.[131]

4.3.5 *Microwave-assisted Knövenagel-Döbner condensation procedures*

The microwave adaptation of the *Döbner* reactions was a promising method, which could shorten the reaction time very effectively.

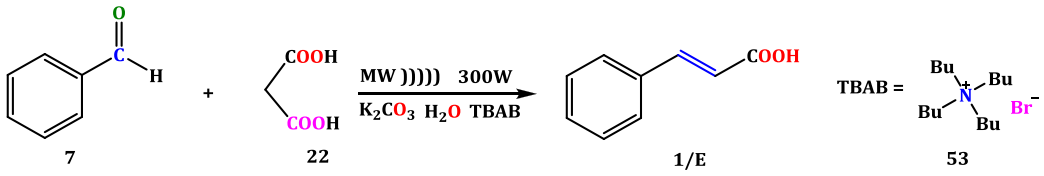


Figure 32 Microwave-assisted *Knövenagel-Döbner* condensation

1 eq. *benzaldehyde* **7**, 1 eq. of *malonic acid* **22**, 0.5 eq. K_2CO_3, 0.5 eq. of *tetrabutyl ammonium bromide* **53** and distilled water (energy transfer material) are placed in a rbf., attached to an air-cooled reflux condenser. The reaction mixture was irradiated with microwave at 300 W for 15 min, using automated pulse sequence and *nitrogen* stream to keep the temperature at max. 90°C. The mixture was left to cool down; then, diluted with 10% KOH. It was extracted with *diethyl ether*, and the combined *aqueous* phases were treated with cc. HCl, until it was strongly acidic to *litmus*. The precipitated crystals (**1/E**) were filtered, washed with *water* and recrystallized from MeOH:H_2O (1:3). The yield varied from 60–87% with 99% purity (1H NMR); thus, the method provided with lower yield than the classic methods, and required expensive phase transfer catalyst.

4.3.6 *Preparation of E-cinnamic acid with the Knövenagel-Döbner condensation*

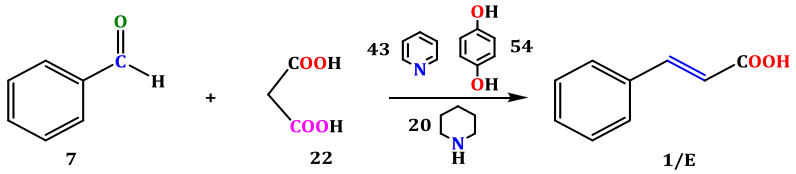


Figure 33 Preparation of *E-cinnamic acid* with *Knövenagel-Döbner* condensation

Malonic acid **22** was placed and dissolved in *pyridine* **43** at 50°C. After all *malonic acid* was dissolved, catalytic amount of *piperidine* **20** was added, and the solution was stirred for 15 min. Traces of *hydroquinone* **54** was also added to the solution. Purified *benzaldehyde* **7** was added dropwise; then, the mixture was stirred for an additional 20 min, and then the temperature was raised to reflux (110°C) for 1–5 hour to complete decarboxylation. The cooled mixture was acidified and the precipitated *E-cinnamic acid* **1/E** was filtered and washed. The crude product was purified by recrystallization from MeOH:H_2O (1:3) and decolorized with activated *charcoal.*[24] Yield: white crystals, 95–100% (99%, ^{1}H NMR). The ^{1}H NMR spectrum is the same as for the other *E-cinnamic acid* **1/E** samples.

4.3.7 *Preparation of E-3-(2 or 3-furyl-, 2 or 3-thienyl-, 2, 3 or 4-pyridyl- and 2-selenophenyl)propenoic acids with the Knövenagel-Döbner condensation*

The preparation of the *E-heteroaryl propenoic acids* are the same as applied in section *4.3.6*. The *pyridyl* derivatives were treated a bit differently, due to the basic characteristics of the *pyridine* ring. *Pyridine* **43** was evaporated on a rotary evaporator and the residue was diluted with *diethyl ether*; then, filtered and washed with ice-cold water. The prepared *heteroaryl* derivatives are shown in *Table 1*. The ^{1}H NMR spectra of the products show full consistency with the structures of each derivative; thus, the syntheses of the desired materials were successful in all cases and each derivative was the pure *E* isomer.

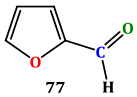
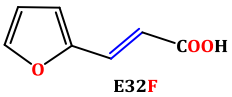
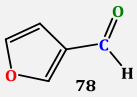
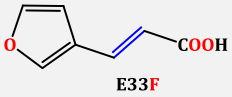
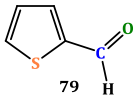
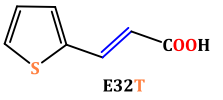
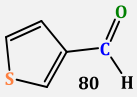
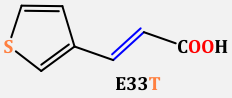
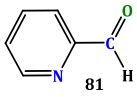
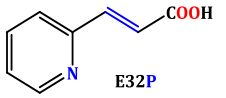
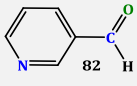
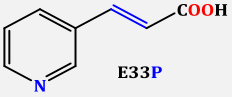
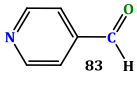
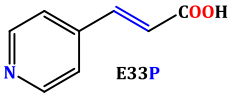
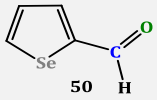
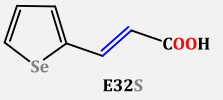
Aldehyde	Reaction time	Yield and purity	Structure of the product	Mr.	mp. (°C)
 77	3 h	97% (99% ^{1}H NMR[132])	 E32F	138.12	140
 78	4 h	98% (99% ^{1}H NMR*)	 E33F	138.12	153
 79	5 h	91% (99% ^{1}H NMR[133])	 E32T	154.19	146
 80	5 h	99% (99% ^{1}H NMR*)	 E33T	154.19	153
 81	7 h	77% (99% ^{1}H NMR*)	 E32P	149.15	203*
 82	8 h	83% (99% ^{1}H NMR[134])	 E33P	149.15	234
 83	9 h	87% (99% ^{1}H NMR[135])	 E33P	149.15	293*
 50	12 h	95% (97% ^{1}H NMR*)	 E32S	201.08	140

Table 1 Synthetic parameters and results for the *E-3-heteroaryl propenoic acid* preparation with *Knövenagel-Döbner* method

4.3.8 *Solubilities of cinnamic acid derivatives in DMSO*

The solubilities of the *E-cinnamic acid* derivatives were determined by gravimetry at 25°C in DMSO (spectroscopic grade) for NMR spectroscopy measurements. The results are listed in *Table 2*.

Compound	Solubility (g/l)	Solubility (M)
E32F	1671 ± 14	12.1 ± 0.1
E33F	1022 ± 12	7.4 ± 0.1
E32T	1526 ± 15	9.9 ± 0.1
E33T	1095 ± 11	7.1 ± 0.1
E32P	177 ± 3	1.19 ± 0.02
E33P	89 ± 1	0.6 ± 0.01
E34P	51 ± 1	0.34 ± 0.01
E32S	123 ± 6	0.61 ± 0.03
E2Ph33P	550 ± 11	2.44 ± 0.05
E23T33T	739 ± 4	3.12 ± 0.02

Table 2 Solubilities in DMSO at 25°C

4.3.9 *Preparation of octyl-E-3-(2-thienyl)propenoate and octyl-E-3-(3-thienyl)propenoate*

To determine the orientation of the *self-assembled* layers, the *octyl esters* of the *thienyl monosubstituted* derivatives were chosen as model compounds, due to the unreactive, long *alkyl-chain*, preventing the formation of *hydrogen-bonded* dimers.

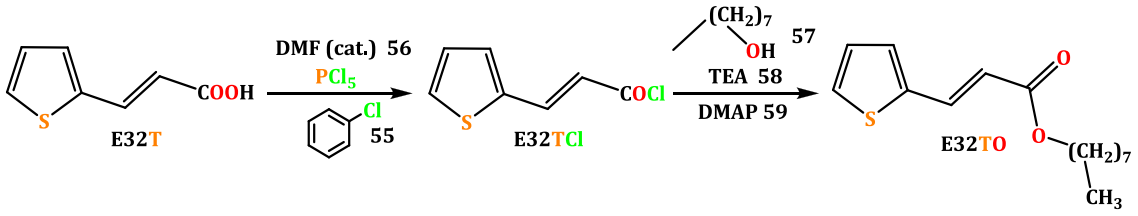


Figure 34 The esterification of **E32T** and **E33T** with *n-octanol* (**57**)

1 eq. of **E32T** or **E33T** was dissolved in **55** and the solution was cooled to 0°C. 1 eq. PCl_5 was added in one portion. The system was equipped with a *water* trap and intensely stirred until the evolution of HCl stopped; then, the temperature was raised to 100°C and was kept there for 2 h. The resulting dark solution was evaporated in vacuo, and the black residue, which was either *E-3-(2-thienyl)propenoyl chloride* **E32TCl** or *E-3-(3-thienyl)propenoyl chloride* **E33TCl**, was used without further purification. 1 eq. *n-octanol* **57** and 4 eq. TEA was dissolved in *dichloromethane* and catalytic amount of DMAP **59** was added, then the mixture cooled down to 0°C. 1 eq. of **E32TCl** or the **E33TCl** was dissolved in *dichloromethane*; then, it was added to the stirred solution dropwise, and the mixture was stirred for an additional 1 h, at rt. The resulting mixture was washed with *brine* and extracted with *dichloromethane* and dried over sicc. $MgSO_4$. The volatiles were evaporated on a rotary evaporator, and the residue was purified by flash column chromatography on *Normasil 60* silica gel: EtOAc-Hexane 1:1 (**E32TO**: R_f = 0.88; **E33TO**: R_f = 0.81). Yield: 90% of the theoretical for both derivatives, (purity: 99% for both derivatives, GC). The structures were verified with ^{1}H NMR spectroscopy. **E32TO: ^{1}H NMR** (500 MHz, $CDCl_3$) δ = 0.86–0.89 (t, J = 7.1 Hz, 3H), 1.27–1.38 (m, 8H), 1.67–1.69 (quint, J = 7.3 Hz, 2H), 4.16–4.19 (t, J = 6.7 Hz, 2H), 6.24–6.27 (d, J = 15.8 Hz, 1H), 7.27–7.28 (d, J = 4.8 Hz, 1H), 7.30–7.32 (dt, J_A = 4.5 Hz, J_B = 2.4 Hz, 1H), 7.47 (d, J =2.3 Hz, 1H), 7.64–7.67 (d, J = 15.8 Hz, 1H). **E33TO: ^{1}H NMR** (500 MHz, $CDCl_3$) δ = 0.87–0.89 (t, J = 7.0 Hz, 3H), 1.28–1.39 (m, 8H), 1.66–1.70 (quint, J =6.7 Hz, 2H),

4.17–4.19 (t, J = 6.7 Hz, 2H), 6.24–6.28 (d, J= 15.6 HZ, 1H), 7.28–7.29 (d, J = 4.9 Hz, 1H), 7.31–7.32 (d, J = 4.7 Hz, 1H), 7.48 (s, 1H), 7.64–7.69 (d, J =15.8 Hz, 1H).

4.4 General methods for the synthesis of *E-Z*-cinnamic acid isomer mixtures

4.4.1 Preparation of E-Z cinnamic acid isomer mixtures via UV induced photoisomerization and the kinetics of the process

The *molar absorptivity constant* (ε ($M^{-1}cm^{-1}$)) determined for each *E-cinnamic acid* and for some of the disubstituted derivatives as well, at the wavelength belongs to the absorbance maximum, at 25°C (298K) in EtOAc (99.9% Sigma-Aldrich, HPLC grade) in a 10mm *quartz* cuvette (Hellma). The proper solutions diluted until their absorptions reach approximately the 0.8-1.3 absorbance value, where the *Beer-Lambert* equation is still valid and linear (*Table 3*). The isomerization experiment was performed with the same equipment in EtOAc, using the 0.05 M solution of the studied isomer. The process was followed with ^{1}H NMR spectroscopy (UV-Vis spectroscopy could not be applied due to the overlapping absorption bands of the isomers). The solutions were irradiated using the UV-reactor setup (*4.2.3*) for different time frames (15 min, 30 min, 1 h, 2 h, 4 h, 8 h, 12 h, 24 h, 48 h) and the isomer distribution was calculated with the relative integration of the peaks of the olefinic protons. The results (which are very close to the equilibrium isomer distribution) are shown in the last column of *Table 3*.

Compound	c(M)	λ_{max} (nm)	ε ($M^{-1}cm^{-1}$)	*Z:E* ratio (24h)	$K_{equilibrium}$
E32F	3.62E-05	298	22652	29:71	0.40
E33F	7.24E-05	271	13149	35:65	0.53
E32T	6.49E-05	305	16884	30:70	0.42
E33T	6.49E-05	276	19783	66:34	1.94
E32P	6.70E-05	287	13926	87:13	6.69
E33P	6.44E-05	279	16935	13:87	0.15
E34P	5.90E-05	257	23220	2:98	0.02
E32S	1.24E-04	318	7890	42:58	0.72
1/E	6.75E-05	272	19556	58:42	1.38
Z32F	7.24E-05	299	14516	N/A	N/A
Z32T	8.11E-05	306	16221	N/A	N/A
E2Ph33P	4.43E-05	260 and 285	10632	8:92	0.08
E23T33T	4.21E-05	287	14180	60:40	1.51

Table 3 Parameters of the isomerization experiments followed by UV-Vis and ^{1}H NMR spectroscopy

4.5 General methods for the stereoselective synthesis of *Z*-cinnamic acids

4.5.1 Preparation of 3-phenylpropiolic acid and methyl 3-phenylpropiolate

3-phenylpropiolic acid **26** and *methyl 3-phenylpropiolate* **60** were used as model compounds to optimize the *Z*-selective hydrogenation with *Lindlar's* catalyst. The selected synthetic role can be seen on *Figure 35*.

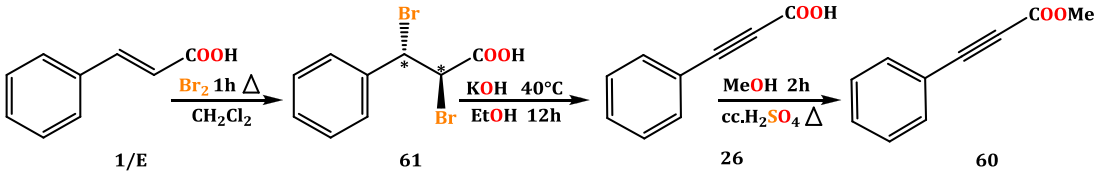


Figure 35 The preparation role of *3-phenylpropiolic acid* (**26**) and its *methyl ester* (**60**) from *E-cinnamic acid* (**1/E**)

Preparation of (2R, 3S)-dibromo-3-phenylpropanoic acid 61: 1 eq. of *E-cinnamic acid* was dissolved in *dichloromethane* and 1.05 eq. *bromine* in *dichloromethane* was added and the mixture was refluxed and stirred for 1 h. The result was cooled to 0°C and the precipitated crystals were filtered off and washed with ice cold *dichloromethane.*[136] The product is a white crystalline material **61**. Yield: 90% of the theoretical (99%, ^{1}H NMR). The structure of **61** was consistent with the spectrum*.

Preparation of 3-phenylpropiolic acid 26: 1 eq. of crude **61** was added to a solution of 4 eq. KOH in 96% *aqueous ethanol*. The solution should be kept under 40°C during the addition. Then, the mixture was stirred overnight (8h) at 40°C. The resulting brown solution was filtered and the filtrate was treated with cc.HCl until it became acidic to litmus; then, the precipitated crystals were filtered off again. The *ethanol* was distilled off the filtrate, and the resulted solution was mixed with the separated salts and the solution was treated with 20% H_2SO_4. The precipitated crystals were filtered off and washed with cold *water*. The crude product was recrystallized from MeOH:H_2O (1:3) and was decolorized with activated *charcoal*. The product is a yellow crystalline material with pleasant smell **26**[137]. Yield: 75% of the theoretical (95%, ^{1}H NMR). The structure was consistent with the spectrum.[138]

Preparation of methyl 3-phenylpropiolate 60: 1 eq. **26** was dissolved in 10 eq. MeOH and catalytic amount of cc.H_2SO_4 was added while the solution was intensely swirled.

The solution was refluxed for 5 hours. The *methanol* was evaporated; then, the residue was washed with 5% Na_2CO_3 solution, and was extracted with *diethyl ether*. The combined organic phases were dried over sicc. $MgSO_4$; then, the *ether* was evaporated. The residue was vacuum-distilled at 100 mbar, and the fraction between 140-160 °C was collected as yellow oil **60**, which had pleasant smell.[139] Yield: 60% of the theoretical (98%, GC).

Preparation of the Lindlar's catalysts: The fundamental *Lindlar's* catalyst (5% *Pd* on $CaCO_3$, conditioned with Pb^{2+} ions) is commercially available; however, the preparation is uncomplicated; therefore, the catalyst was prepared using two different catalyst supports ($CaCO_3$ and $BaSO_4$).

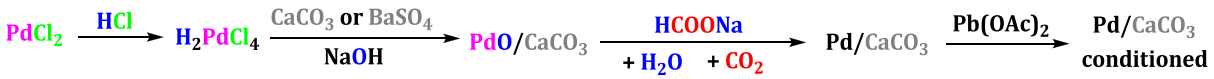


Figure 36 The preparation of conditioned *Lindlar's* catalyst from *palladium(II) chloride*

$PdCl_2$ was dissolved in cc.HCl. The resulting H_2PdCl_4 solution was diluted with *water* and the pH was brought to 4.0–4.5 by 3M NaOH. Precipitated *calcium carbonate* (or $BaSO_4$) was added to the solution. The well-stirred suspensions were heated to 75–85°C and were held at this temperature until all the *palladium* was precipitated. HCOONa solution (about 0.7M) was added, and the reduction was completed by stirring the mixtures at 75–85°C for 40 min. The catalysts, which were black now, were separated and were washed multiple times with *water*. The wet catalysts were suspended in *water* again and 0.25 eq. 20% $Pb(OAc)_2$ solution was added. The slurries were stirred and heated at 75–85°C for 45 min. The catalysts were separated and washed with *water*, and dried in an oven at 60–70°C. The dried catalysts, were dark grey powders (5% *Pd* on $CaCO_3$ and $BaSO_4$).[140]

4.5.2 Z-selective hydrogenation of 3-phenylpropiolic acid with Lindlar's catalyst

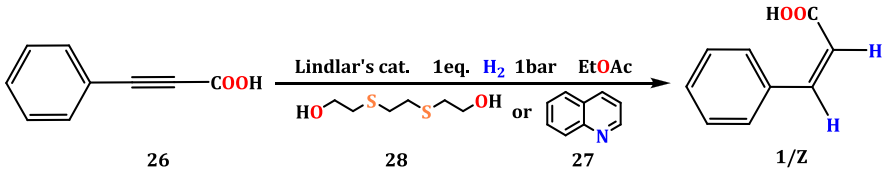


Figure 37 *Z*-selective hydrogenation of (**26**) with *Lindlar's* catalyst under different conditions

1 eq. of *3-phenylpropiolic acid* **26** was dissolved in EtOAc; then, was placed in a hydrogenation apparatus (*Figure 38*). *Lindlar's* catalyst (5%, *Pd*/$CaCO_3$ or *Pd*/$BaSO_4$) and 0.1 eq. *quinoline* **27** or *3,6-dithia-1,8-octanediol* **28** were added to the mixture. The apparatus was evacuated; then, was refilled with H_2 (1 bar); then, stirring was started. If the first equivalent of H_2 absorbed the reaction was stopped to avoid perhydrogenation. The catalyst was filtered off, and the residue was concentrated in vacuo; then, was diluted with *diethyl ether*, and washed with 10% HCl solution. The organic phase was separated; then, dried over sicc. $MgSO_4$, and the ether was removed on a rotary evaporator.[141] The residue was a brown oily mixture, containing several derivatives. The results are listed in *Table 4*.

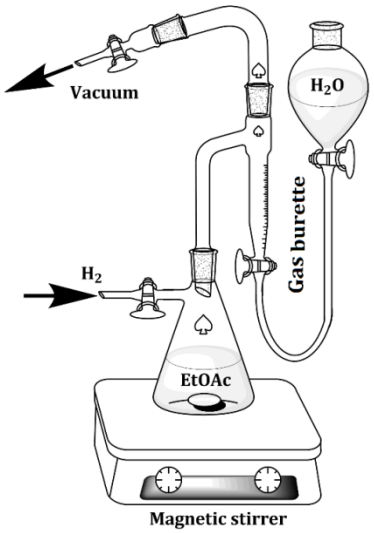


Figure 38 The hydrogenation apparatus

Catalyst	Poison	Yield	*Z:E* ratio
Pd/$CaCO_3$ (Pb^{2+})	**27**	85–95%	90:10
Pd/$CaCO_3$ (Pb^{2+})	**28**	80–92%	95:5
Pd/$BaSO_4$ (Pb^{2+})	**27**	85%	92:8
Pd/$BaSO_4$ (Pb^{2+})	**28**	87%	94:6

Table 4 Results for the different hydrogenation variants

4.5.3 *Utilizing the HWE synthesis for Z-selective preparation*

4.5.4 *Preparation of the different HWE phosphonates*

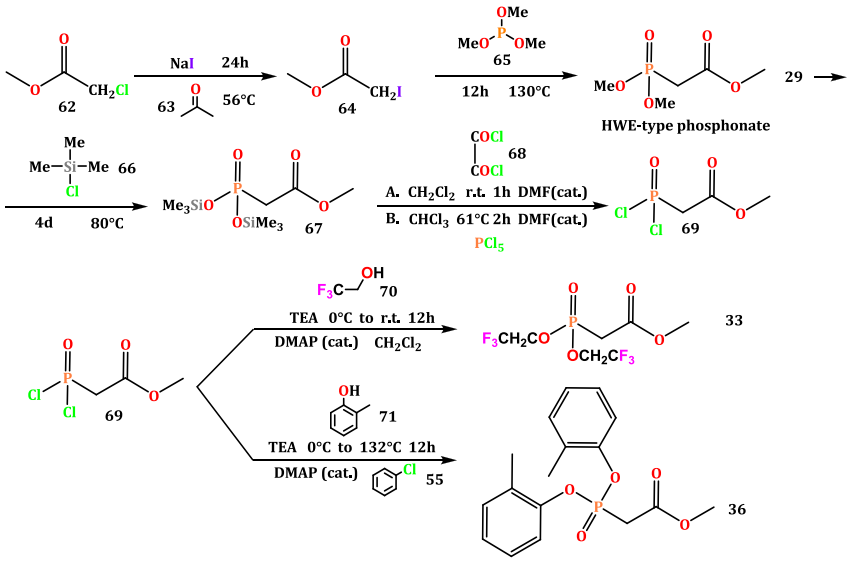


Figure 39 Reaction role to the *Z*-selective reagents (**33**,**36**)

Preparation of methyl iodoacetate: 1 eq. of **62** was dissolved in *acetone* **63** and 1 eq. sicc. NaI was added. The mixture was vigorously stirred and refluxed for 24 h. After the appropriate time elapsed, the precipitated salts were filtered off, and the filtration was evaporated. The dark brown residue was washed with 5% $Na_2S_2O_3$ solution, and then extracted with *diethyl ether*. The combined organic phases were dried over sicc. $MgSO_4$; then, the volatiles were removed. The residue vacuum distilled at 100 mbar and the fraction between 60–65°C was collected as a colorless, highly irritative liquid, which is **64**.[142] Yield: 80% of the theoretical (96%, GC).

Preparation of trimethyl phosphonoacetate 29 with Michaelis-Arbuzov reaction: 1 eq. **64** and 2 eq. *trimethyl phosphite* **65** were refluxed at 110-130°C for 12h. The excess of **65** was removed; then, the residue was vacuum-distilled at 100 mbar, and the fraction between 120–128°C was collected as a colorless oily compound, which is **29**.[143] Yield: 91% of the theoretical (93%, GC).

Preparation of methyl bis(trimethylsilyl)phosphonoacetate 67: 1 eq. of **29** and 5 eq. *trimethylsilyl chloride* **66** were placed in a screw-cap *Pyrex* bottle, equipped with a pressure meter. The temperature was raised slowly to 80°C, while the pressure was monitored and kept under 3 bars to avoid explosion. When the temperature and the

pressure were stabilized, the mixture was kept at 80°C for 4 days. The volatiles were removed under reduced pressure, and the residue was vacuum-distilled at 100 mbar. The fraction between 142-144°C was collected as colorless oil, which was pure **67**.[144] Yield: 95% of the theoretical (93%, GC).

Preparation of methyl dichloro phosphonoacetate 69: Method A: 1 eq. of **67** was dissolved in CH_2Cl_2 and one drop of DMF was added. The system was flushed with *argon*; then, was cooled to 0°C. 2.5 eq. of *oxalyl chloride* **68** was dissolved in CH_2Cl_2; then, was added dropwise to the stirred solution. The ice bath was removed, and the solution was stirred for an additional hour. The volatiles were removed under reduced pressure, the residue, which was a yellow oil **69** used in the next step without further purification.[144] Yield: 96% of the theoretical (91%, ^{31}P NMR*). **Method B.**: 1 eq. **67** was dissolved in *chloroform* and one drop of DMF was added. 1 eq. of PCl_5 was suspended in *chloroform* and was added dropwise to the cooled solution under vigorous stirring. The temperature was elevated to 60°C, and was kept for 2 hours. The mixture was cooled and the volatiles were removed under reduced pressure. The residue was a reddish oil **69**, which was used without further purification in the following step.[145] Yield: 97% of the theoretical (90%, ^{31}P NMR*).

Synthesis of the Still-Gennari phosphonate 33: A mixture of 6 eq. TEA and 2.5 eq. *trifluorethanol* **70** was dissolved in CH_2Cl_2 and was cooled to 0°C, and then a small crystal of DMAP was added. The flask was flushed with *argon*, and a mixture of 1eq. **69** in *dichloromethane* was added dropwise to the cooled solution under vigorous stirring. The ice bath was removed, and the solution was stirred for additional 12 h. The resulted mixture was washed with *brine*, and was extracted with CH_2Cl_2. The organic layers were combined and dried over sicc. $MgSO_4$, and the volatiles were evaporated in vacuo. The residue, which was a yellow oily material **33**, was purified via flash column chromatography on *Normasil 60 silica gel*, eluted with *n-hexane*-EtOAc, 2:1. R_f = 0.61.[144] Yield: 87% of the theoretical (95%, ^{31}P NMR*).

Synthesis of Ando-type phosphonate 36: 2.5 eq. of *o-cresol* **71** and 6 eq. of TEA were dissolved in **55** and a catalytic amount of DMAP was added as well. The flask was flushed with *argon*, and was cooled to 0°C. 1 eq. of **69** dissolved in **55**, and then added dropwise to the well-stirred solution. After the addition the temperature raised to reflux for additional 12h. The further treatment was the same as was used in the previous

method. The residue, which was a dark brown oily material **36** was purified via flash column chromatography on *Normasil 60 silica gel*, eluted with *n-hexane*-EtOAc, 2:1. R_f=0.70. Yield: 40% of the theoretical (93%, ^{31}P NMR*).

4.5.5 *Optimization of the Z-selective HWE reaction with Still-Gennari and Ando-type phosphonates*

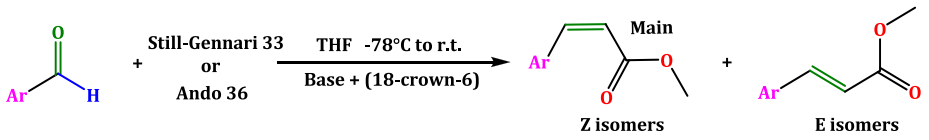


Figure 40 Optimization of the HWE reaction to synthesize *Z-cinnamic acids*

The general conditions for the optimization: the solvent was water-free THF in each case, combined with 1 eq. of the proper base (NaH, *t*-BuOK **72**, KHMDS **32**) and sometimes with an additive (*18-crown-6* **31**), 1 eq. as well. This mixture was injected under vigorous stirring into the cryogenic apparatus (*Figure 41*), which was flushed with *argon*, and was cooled from 0 to −78°C. After 15 min stirring, 1 eq. of the proper *phosphonate* was injected. After another 15 min stirring, 1 eq. of the *aromatic aldehyde* (which was **7** during the optimization) was injected. The mixture was stirred from 5 min to 1 h at −78°C to 0°C or to r.t. for 12 h; then, it was poured into 10% NH_4Cl solution to quench the reaction. The residue was washed with *brine*, and was extracted with EtOAc. The combined organic layers were dried over sicc. $MgSO_4$, and the volatiles were removed on a rotary evaporator. The residue was analyzed with GC and ^{1}H NMR to determine the yield and the isomer distribution. The results are listed in *Table 5*.

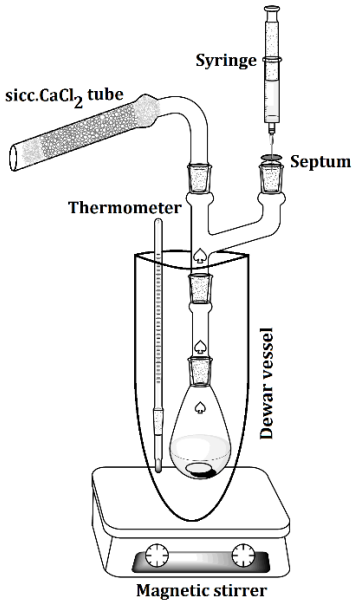


Figure 41 The cryogenic apparatus applied for the *Z*-selective HWE synthesis

Entry	Phosphonate	Base	Additive	Conditions	*Z:E* ratio	Yield
1	Ando **36**	KHMDS	18-crown-6	-78°C 1h -> 12h r.t.	96:4	20%
2	Ando **36**	KHMDS	-	-78°C, 1h	76:24	75%
3	Ando **36**	*t*-BuOK	18-crown-6	-78°C, 1h	67:33	88%
4	Ando **36**	*t*-BuOK	-	-78°C, 1h	90:10	80%
5	Ando **36**	NaH	18-crown-6	-78°C 5min -> 0°C 30min	73:27	9%
6	Ando **36**	NaH		-78°C 5min -> 0°C 30min	66:34	97%
7	S.-G. **33**	KHMDS	18-crown-6	-78°C, 1h -> 12h r.t.	77:23	90%
8	S.-G. **33**	KHMDS	-	-78°C, 1h	84:16	85%
9	S.-G. **33**	*t*-BuOK	18-crown-6	-78°C, 1h	76:24	87%
10	S.-G. **33**	*t*-BuOK	-	-78°C, 1h	68:32	74%
11	S.-G. **33**	NaH	18-crown-6	-78°C 5min -> 0°C 30min	34:66	85%
12	S.-G. **33**	NaH	-	-78°C 5min -> 0°C 30min	73:27	88%
13	HWE **29**	KHMDS	18-crown-6	-78°C, 1h	34:66	91%
14	**67**	KHMDS	18-crown-6	-78°C, 1h	35:65	7%

Table 5 Results of the optimization of the *Z*-selective HWE synthesis

The *Ando* reagent **36**, shows the best results, when *t*-BuOK **72** applied as a base without any additives (Entry: 4 on *Table 5*); therefore, this variant was chosen for the synthesis of the *Z-heteroaryl cinnamic acid methyl esters*.

The successful reaction mixtures, worked up with flash column chromatography on *Normasil 60 silica gel*, were eluted with MeOH-$CHCl_3$, 5:95. The purified *methyl esters* were hydrolyzed in THF/4 eq. KOH solution at r.t. for 12 h. The resulting mixture was neutralized with cc. HCl. The precipitated crystals were filtered, washed with water and then dried. The *pyridyl derivatives* form *hydrochloride salts* so they should be titrated with *piperidine* **20** to get the precipitated crystals. The applied reaction conditions and the chromatography parameters for the *Z-methyl heteroaryl cinnamates* are shown in *Table 6*.

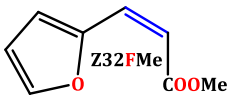
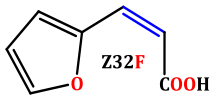
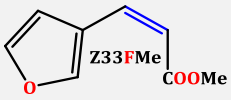
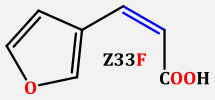
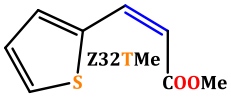
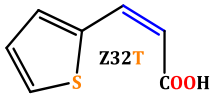
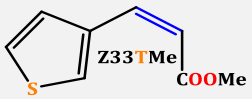
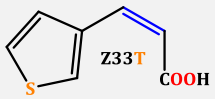
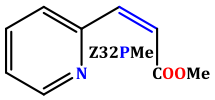
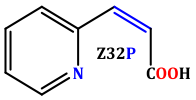
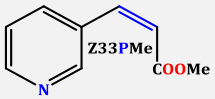
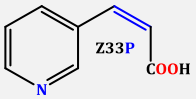
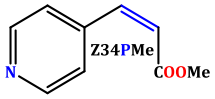
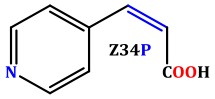
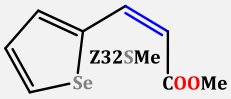
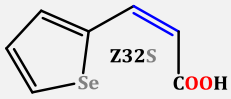
***Z:E* ratio**	**Yield (GC):**	**Structure of the Me-ester products**	**R_f of the Me-esters**	**Structure of the final products**
87:13	87%		0.78	
97:3	94%		0.83	
89:11	70%		0.75	
87:13	75%		0.76	
84:16	97%		0.29	
89:11	90%		0.32	
93:7	86%		0.31	
90:10	60%		0.82	

Table 6 The results of the *Z*-selective synthesis *of heteroaryl cinnamyl esters* with *Ando* reagent (**36**)

The structures of the *Z-methyl heteroaryl cinnamates* were determined by ^1H NMR spectroscopy, and it was found that each derivative was successfully synthesized, both in the *methyl ester* and the *acidic forms*.

The ^{1}H NMR spectra of the *Z-methyl heteroaryl cinnamates* and the acids – they are not available in the literature – were the following:

The ^{1}H NMR spectra of the *Z-methyl heteroaryl cinnamates*:

Z33F–Me: ^{1}H NMR (500 MHz, $CDCl_3$) δ = 3.76 (s, 3H), 5.78–5.81 (d, J = 12.5 Hz, 1H), 6.69–6.71 (d, J = 12.5 Hz, 1H), 6.92 (s, 1H), 7.39 (s, 1H), 8.11 (s, 1H).

Z34P–Me: ^{1}H NMR (500 MHz, $CDCl_3$) δ = 3.69 (s, 3H), 6.11–6.13 (d, J = 12.5 Hz, 1H), 6.86–6.89 (d, J = 12.5 Hz, 1H), 7.36 (br, 2H), 8.59 (br, 2H).

Z32S–Me: ^{1}H NMR (500 MHz, $CDCl_3$) δ = 3.78 (s, 3H), 5.74–5.76 (d, J = 12.2 Hz, 1H), 6.96–6.99 (d, J = 12.4 Hz, 1H), 7.26-7.28 (d, J = 12.5 Hz, 1H), 7.42 (br, 1H), 7.54 (br, 1H), 8.31 (br, 1H).

The ^{1}H NMR spectral data of the *Z-heteroaryl cinnamic acids*:

Z33F: ^{1}H NMR (500 MHz, DMSO-d_6) δ = 5.74–5.77 (d, J = 12.5 Hz. 1H), 6.77–6.79 (d, 12.5 Hz, 1H), 7.02 (br, 1H), 7.67 (br, 1H), 8.22 (s, 1H).

Z34P: ^{1}H NMR (500 MHz, MeOD-d_4) δ = 6.15–6.17 (d, J = 12.2 Hz, 1H), 6.58–6.61 (d, J = 12.3 Hz, 1H), 7.42 (s, 2H), 8.55 (s, 2H).

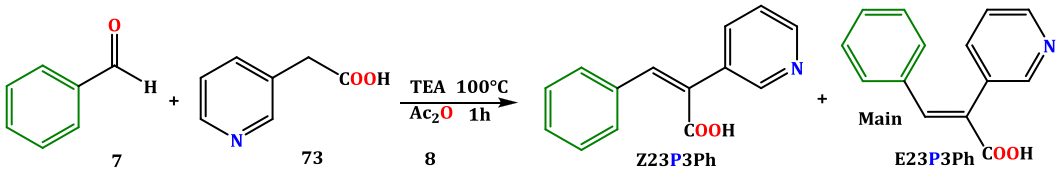
Z32S: ^{1}H NMR (500 MHz, DMSO-d_6) δ = 5.68-5.70 (d, J = 12.2 Hz, 1H), 7.28-7.30 (d, J = 11.8 Hz, 1H), 7.37–7.38 (d, J = 3.9 Hz, 1H), 7.59–7.60 (d, J = 3.3 Hz, 1H), 8.24–8.25 (d, J= 5.4 Hz, 1H).

4.6 Synthesis of the α-substituted cinnamic acid derivatives

Due to the synthetic work of the former members of our research group, several *α-substituted cinnamic acids* were prepared via the modified *Perkin* condensation. Thus, many *thienyl-substituted* variants were available at the beginning of my work. They have already been shown in *Figure 24* (each possible variations and isomers with one or two *thienyl* rings). The *furyl* derivatives were widely studied materials, therefore, this part of the compound family was skipped, and we tried to synthesize the *pyridyl* derivatives. As it has been mentioned before, the synthesis of substituted *selenophenyl* compounds were not done (only the *E-Z-3-(2-selenophenyl)propenoic acids* were synthesized), due to safety considerations. The modified *Perkin* condensation between the proper *heteroaryl carboxaldehyde* and the *heteroaryl acetic acid* in *acetic anhydride* **8** in the presence of

TEA, as it described in the literature part, was applied for the synthesis of the *pyridyl* derivatives, the modified *Döbner* condensation and the HWE olefination method were also tried, and this latter was found to be the best synthetic route to *α-substituted pyridylcinnamic acids*. The general procedures applied for the synthesis are listed in the followings.

4.6.1 The synthesis of E-Z-3-phenyl-2-(3-pyridyl)propenoic acid via the modified Perkin condensation



7 73 8 Z23P3Ph E23P3Ph

Figure 42 The synthesis of **Z3Ph23P** with the modified *Perkin* condensation

1.1 eq. of *3-pyridylacetic acid* **73** and 1 eq. of **7** were dissolved in the mixture of TEA and Ac_2O, 2:1. The mixture was brought to 100°C on a water bath and was kept at this temperature for 1 h. The dark reaction mixture was evaporated, and the residue was treated with cc. HCl; then, the unreacted *aldehyde* was removed with extraction (*diethyl ether*). The acidic mixture was titrated with piperidine **20** at 0°C until the product started to precipitate. The precipitate was a mixture of *pyridyl* compounds to be further processed via column chromatography on *Normasil 60* eluted with DMSO. Yield: 5% of the theoretical, isomer distribution: 72:28; E:Z (^{1}H NMR*).

Unfortunately, this method provided with low yield, and the separation of the resulting materials were difficult (because the title material only soluble in DMSO which is a too powerful eluent); nevertheless, both isomers could be synthesized via the reaction. The preparation of these compounds was tried with the *Döbner* method, but the reaction did not take place (yield: <1%).

4.6.2 The synthesis of E-2-phenyl-3-(3-pyridyl)propenoic acid via the HWE method

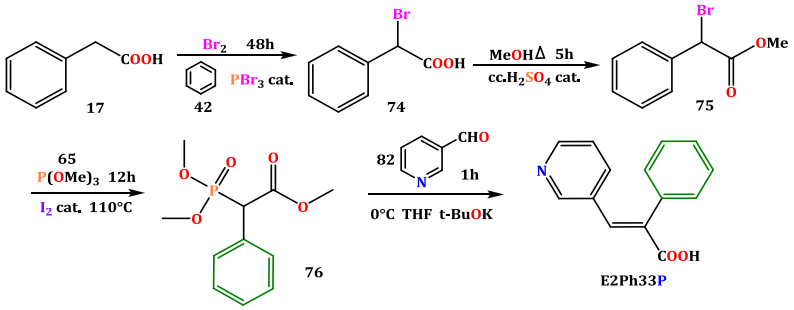


Figure 43 Preparation of *trimethyl phosphono-2-phenylacetate* and the condensation of it with *3-pyridinecarboxaldehyde*

Preparation of 2-bromophenylacetic acid 74: 1 eq. of *phenylacetic acid* **17** was dissolved in *benzene* **42** and 1.1 eq. of *bromine* was added. Catalytic amount of PBr_3 was added as well, and the mixture was gently refluxed for 48 h. The *benzene* was removed via distillation; then, the residue was dissolved in *petroleum ether* and was stored at −20°C for 12 h. The precipitated crystals **74** were filtered off and dried.[146] Yield: 65% white crystalline product (97% ^{1}H NMR*).

Preparation of methyl 2-bromo-2-phenylacetate 75: 1 eq. of **74** dissolved in 15 eq. of anhydrous *methanol* and catalytic amount of cc. H_2SO_4 was added. The mixture was refluxed for 5 h, and the excess *methanol* was removed under reduced pressure. The residue was washed with saturated Na_2CO_3 solution, and was extracted with *diethyl ether*. The ether was removed, and the residue was distilled at 100 mbar and the fraction between 160−165°C was collected as colorless oil **75**. Yield: 68% of the theoretical (98%, GC).

Preparation of trimethyl 2-phenyl phosphonoacetate 76 (Michaelis-Arbuzov): 1 eq. of **75** was dissolved in 4 eq. of *trimethyl phosphite* **65**, and a small crystal of *iodine* was added. The mixture was refluxed for 12 h; then, the volatiles were removed under reduced pressure. The residue was a dark brown oily compound **76**, 67% of the theoretical (95%, GC).

Preparation of E-2-phenyl-3-(3-pyridyl)propenoic acid: 1 eq. of *t*-BuOK was dissolved in THF, and was cooled to 0°C for 15 min. 1 eq. of **76** was added (as a THF solution), and the residue was stirred for an additional 15 min. 1 eq. of *3-pyridinecarboxaldehyde* **82** was injected, and the solution stirred for 1 h at 0°C. The

reaction was quenched with NH_4Cl solution; then, extracted with EtOAc. The solvent was removed, and the resulting *methyl ester* (65% of the theoretical) was stirred with 4 eq. KOH in THF for 12h at r.t. The THF was evaporated, and then the residue treated with cc. HCl until it's became strongly acidic. The mixture cooled to r.t. and titrated with piperidine **20** until the formation of the precipitate stopped. The precipitated crystals were filtered off, washed with water and dried. Yield: ~100% quant. in two steps (99% ^{1}H NMR, 100% *E*). **E2Ph33P-Na: ^{1}H NMR** (500 MHz, D_2O) δ= 6.68-6.70 (dd, J_A= 3.8 Hz, J_B= 7.6 Hz, 1H), 6.77-6.78 (m, 2H), 6.86-6.87 (d, J= 7.9 Hz, 1H), 7.03-7.04 (m, 3H), 7.08 (s, 1H), 7.82 (s, 1H), 7.86-7.87 (d, J= 3.5 Hz, 1H).

4.7 Computational chemistry part: software and methods

4.7.1 Official software packages

- Each molecular model structure was created with **Arguslab 4.02**[147], which is a bit outdated, shareware, but it is a still reliable quantum chemical tool for building molecular structures fast and easy.
- The preoptimization of the drawn molecules was performed with the **Hyperchem 8.0.6 Pro**.[148] package at low level of theory (**PM3**[149], **HF**[150]**/STO-3G**[151]). The software package is very user friendly; however, its performance is poor (the academic version is available only for 1 CPU core), and there are some problems with the reproducibility of the results generated.
- To use high-level methods and to calculate structures and energies precisely for larger systems as well, the **Gaussian 09' rev. D.01**[152] was used. The version is not locked to 1 CPU core; therefore, it performs much better with the relevant hardware and provides with reproducible data.
    - To extract useful data from the *Gaussian* output files, the **Avogadro v1.1.1**[153] program was applied, which is a free cross-platform molecular editor and visualizer.

4.7.2 Creating software for conformational analysis (Quantum Hyppo and Izzy Reader)

Hyperchem 8.0.6 has an automatic conformational analysis module; however, the performance of the program is low, as it was mentioned before, and the energy limit defining stable conformers should be set at the beginning of the calculation, which leads to many unstable conformers. **Gaussian 09'** does not have an automatic tool to perform

conformational analysis; however, its input, output and batch file structures are well-defined, thus, the generation of files and getting out data in very high numbers are possible with some simple scripts, therefore it was decided to write a conformational analysis tool, which can be used with **Gaussian 09'** to setup and perform complex conformational search with a few mouse clicks. The main advantage of this method is its versatility. One does not have to define the terms before the calculations; they always can be modified in the knowledge of the initial results.

The choice was **AutoHotkey v1.1**[153] (AHK, 1.1.16.05, 2014), which is a scripting and automation software for *Windows* operating systems using a simple language-based on **AutoIt**. The programs written are called **Quantum Hyppo v.1.0** (demo version), and **Izzy Reader v1.0** (demo version as well), they are available as downloads from http://www.staff.u-szeged.hu/~csankok/7542/QH_win32.zip free of charge. The current demo versions work properly under: *Vista*, *Win7* and *Win8(.1).* The programs use the following algorithms:

1. Finding the line(s) in the **Gaussian** input file, containing the atomic numbers, which define the dihedral angles. These lines can be found after the 4th empty line (the following example shows the representative file structure of **Gaussian 03-09** .gjf files):

```
%mem=32GB                                                  //define the maximum memory consumption
%nprocshared=16                                            //define the number of CPUs to utilize
%chk=C:\G09\CHK\ethane.chk                                 //place of the checkpoint file
# opt=modredundant freq hf/6-31g(d,p) geom=connectivity    //calculation type and keywords
                                                           //1. empty line
Ethane conformer                                           //title
                                                           //2. empty line
0 1                                                        //charge and spin multiplicity
 C    -4.05770000    0.49200000    0.00000000            //atom labels and coordinates in Å
 C    -2.53940000    0.44160000    0.00000000
 H    -4.40000000    1.55160000    0.00000000
 H    -4.45220000   -0.01900000    0.90730000
 H    -4.45220000   -0.01900000   -0.90730000
 H    -2.19710000   -0.61800000    0.00000000
 H    -2.14490000    0.95260000    0.90730000
 H    -2.14490000    0.95260000   -0.90730000
                                                 //3. empty line
 1 2 1.0 3 1.0 4 1.0 5 1.0                       //connection matrix
 2 6 1.0 7 1.0 8 1.0
 3
 4
 5
 6
 7
 8
                                                 //4. Empty line (this is what we looking for)
D 3 1 2 7 60.000000                              //atom numbers that define a dihedral angle
                                                 //end of input file
```

2. The lines, containing the dihedral data set are counted; then, the dihedral angles, which are floating point values, are changed randomly between the defined angle range

(default: `min= -180.0` to `max= 180.0`). The new file saved uses the following name format: *00001yourfilename.gjf* and the number will be increased by 1 in each cycle. The number of the iteration can be defined as well (`vari=`). After creating the files, the batch file will be created as well (the file routes for inputs and outputs can be defined freely). The graphical user interface of the program is shown in *Figure 44*.

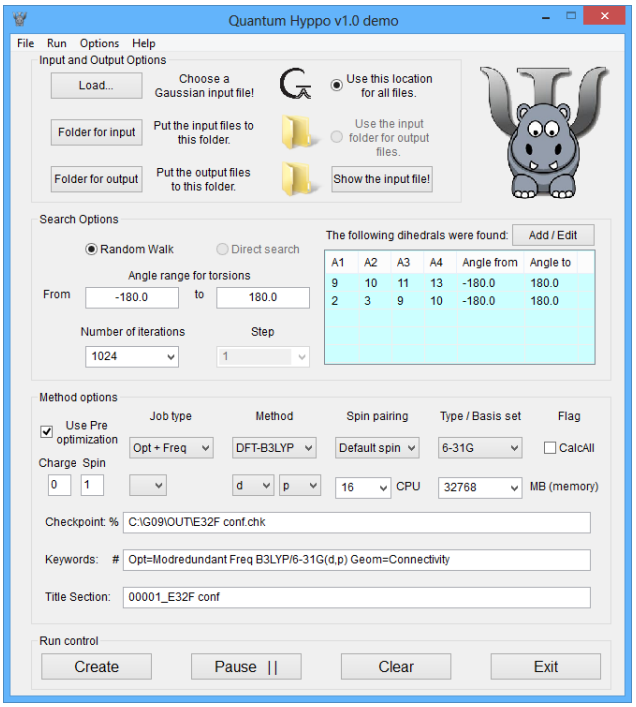


Figure 44 The graphical user interface (GUI) of **Quantum Hyppo** v.1.0 demo (under *Win8*)

3. The main features of the program are as follows: it works with any pregenerated **Gaussian** job file (.gjf); it reads out many data from the file, thus, one can change them easily, including the manipulation with the dihedral angles, the keywords, the specifications and file routes; the created batch file can be executed automatically if one has valid **Gaussian 03'** or **09'** installed on your computer. The simplified source code of **Quantum Hyppo v1.0** demo can be found in the *Appendix.*
4. After getting the output files from the *Gaussian* package, one needs to process these files to get out the useful data like energies, lowest vibrational frequencies, dihedral angles and optimized geometries if they needed. The output files generated by the package are long and complex (4000–12000 lines), thus, no example is presented here, only the relevant parts of the file are extracted like the *zero-point energy*, the *lowest vibrational frequency*, the *dihedral section* and the *optimized geometry* (the following example shows the representative file structure of **Gaussian 03-09'** .out and .log files):

```
|-------------------------------------------------------------------| //section skipped

The following ModRedundant input section has been read:    //the original dihedral angles
 D    9   10   11   13 168.55 F                             //two dihedral angles were found
 D    2    3    9   10 -30.09 F
Iteration  1 RMS(Cart)=  0.02414197 RMS(Int)=  0.34470183
Iteration  2 RMS(Cart)=  0.01171393 RMS(Int)=  0.33740148
|-------------------------------------------------------------------| //section skipped
                         Input orientation:              //the last optimized geometry
 ---------------------------------------------------------------------
 Center     Atomic      Atomic             Coordinates (Angstroms)
 Number     Number       Type              X           Y           Z
 ---------------------------------------------------------------------
      1          6           0       -2.537275    0.838267  -14.237952
      2          6           0       -1.181350    1.210317  -14.523353
      3          6           0       -0.406541    0.354658  -13.829903
      4          8           0       -1.178896   -0.495898  -13.121664
      5          6           0       -2.457884   -0.195477  -13.387693
      6          1           0       -3.430685    1.290555  -14.614791
      7          1           0       -0.843448    2.011196  -15.147645
      8          1           0       -3.196499   -0.798038  -12.905215
      9          6           0        1.030962    0.226111  -13.635736
     10          6           0        1.945301    0.584246  -14.526563
     11          6           0        3.419548    0.482380  -14.349136
     12          8           0        3.891567    0.216787  -13.130259
     13          8           0        4.160915    0.634116  -15.259652
     14          1           0        1.304464   -0.240817  -12.703455
     15          1           0        1.662096    0.966546  -15.489371
     16          1           0        3.236425    0.299252  -12.458212
 ---------------------------------------------------------------------
|-------------------------------------------------------------------| //section skipped
Full mass-weighted force constant matrix:
 Low frequencies ---   -8.7129   -4.3397   -0.0016    0.0051    0.0116    9.1503
 Low frequencies ---   17.5136   54.6257  101.9760  //the lowest frequencies founded
 Diagonal vibrational polarizability:
       69.7748699      44.1963358      91.4472319
 Diagonal vibrational hyperpolarizability:
     2766.4424777    1042.1607792    1897.0719917
 Harmonic frequencies (cm**-1), IR intensities (KM/Mole), Raman scattering
|-------------------------------------------------------------------| //section skipped
                                                  //the energies and thermal corrections
 Zero-point correction=                           0.127830 (Hartree/Particle)
 Thermal correction to Energy=                    0.136236
 Thermal correction to Enthalpy=                  0.137180
 Thermal correction to Gibbs Free Energy=         0.092754
 Sum of electronic and zero-point Energies=           -493.009294
 Sum of electronic and thermal Energies=              -493.000888
 Sum of electronic and thermal Enthalpies=            -492.999944
 Sum of electronic and thermal Free Energies=         -493.044371
|-------------------------------------------------------------------| //section skipped
ERROR IS THE FORCE THAT WELDS MEN TOGETHER....
 TRUTH IS COMMUNICATED TO MEN ONLY BY DEEDS OF TRUTH.      //quote
          TOLSTOI,MY RELIGION
 Job cpu time:  0 days  0 hours  3 minutes  7.0 seconds. //the calculation's time consumption
 File lengths (MBytes):  RWF=     63 Int=      0 D2E=      0 Chk=      4 Scr=      1
 Normal termination of Gaussian 09 at Thu Jan 30 19:05:39 2014.
                                                  //sign of the calculations is converged
```

5. The GUI of the data extracting software is shown in *Figure 45*.

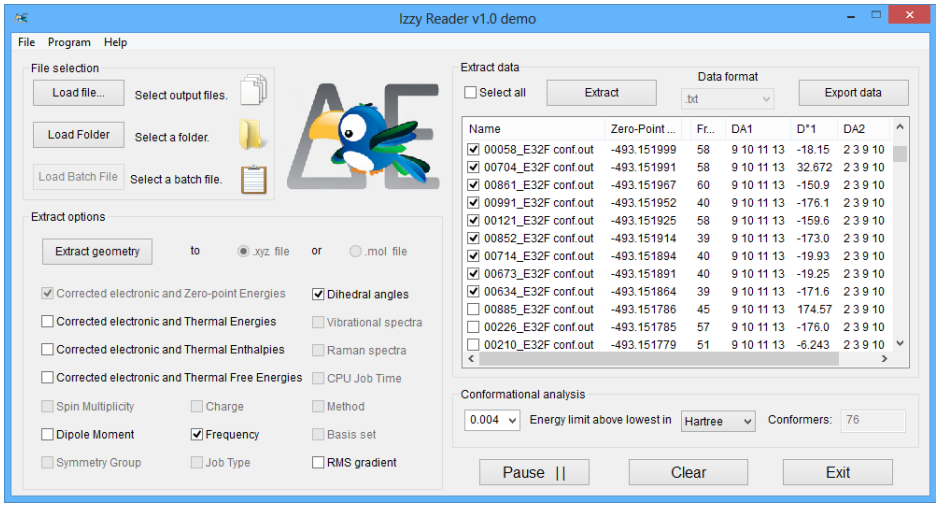


Figure 45 The graphical user interface of **Izzy Reader** v.1.0 demo (under *Win8*)

The main features of the program are as follows: the different *energies*, the *dipole moment*, the *lowest frequency*, the *dihedral sets* and the RMS gradient can be extracted; the list made in a list view can be cut short freely, and can be filtered by the energy to find the conformers; these can be checked by one click and can exported to tab-separated ASCII file; the geometries from the selected files can be extracted as well in .xyz (that giving the geometry by *Cartesian* coordinates and atom symbols in a simple text file).

The simplified source code of **Izzy Reader v.1.0** can be found in the *Appendix*.

4.7.3 The applied quantum chemical methods

As it has been mentioned before, the molecular model of the studied molecules are drawn in **Arguslab 4.02** and are stored in **MDL** .mol files; then, the structures were preoptimized at the **PM3** level (or at **HF/STO-3G** if it was necessary) using **Hyperchem 8.0.6**. The optimized individual monomers (**M**) were translated to **Gaussian 09'** input format, and the two possible dihedral angle sets (determined by the serial number of the 4 – 4 atoms sets) were determined.

- **Conformational analysis of the monomers (M)**

The input files were used to generate 1024 variations for the molecules, which have two variable dihedrals (*cinnamic acids*) and 2048 variations for the *α-substituted cinnamic acids*, due to the three possible rotations, from −180° to 180°, applying the

"*random walk*" method with **Quantum Hyppo v.1.0**. The generated geometries were preoptimized at the **PM6**[154] level with **Gaussian 09'** to find faulty geometries. These were removed from the batch file, then, the rest were optimized at the **HF/6-31G(d,p)**[155] level, and the results were processed to determine the lowest point of the potential energy surface. It belongs to the most stable conformer, which was used as a starting point in further calculations.

- **The geometries and the energies of the monomers (M)**

  From each conformational search, the lowest-energy monomer was selected, and it was further optimized at several different levels like **B3LYP**[156]**/6-31G(d,p)**, **B3LYP/6-31++G(d,p)**[157], **B3LYP/CBSB7**[158], **(MP2**[159]**/6-31G(d,p)), WB97XD**[160]**/6-31++G(d,p)** and **WB97XD/CBSB7** to make references for the *hydrogen-bond* and *π-stacking* energy calculations (**CBSB7** uses the **6-311G(2d,d,p)** basis set with three polarization functions, and it adapts the functions only when needed, e.g. the **2d** function for **S** and **Se** atoms).

- **Geometry and the H-bond energy of the dimers (D)**

  From the refined monomers, dimers (**D**) were built having the *carboxylic* groups close to each other. The dimers were optimized using the above-listed methods (including the **HF/6-31G(d,p)** method and the **B3LYP/CBSB7** method). To determine the correct *H-bond* energy, the calculations were performed using the *Counterpoise* method, (**CP**)[161], which corrects the energy influence coming from the so-called *Basis Set Superposition Error* (**BSSE**).[162]

- **Geometry and the H-bond energy of the tetramers (T)**

  The tetramers are the dimers of the dimers, kept together via weaker second order *hydrogen-bonds* by the *heteroaromatic* rings, via the polarized aromatic *protons* and the *heteroatoms*: C–H…X (X= O, N, S, Se). The calculations were performed at the same levels as for the dimers, and the necessary corrections for the energy calculations have also been made. Due to the higher complexity of the *α-susbtituted cinnamic acids* the calculations for their tetramers are skipped.

- **Geometries and bond energies for the various π-stacked dimers**

  The stacked dimers were constructed from the optimized monomers making the "*sandwich*" form, since this the most favored geometry. The calculations were performed by the **WB97XD/6-31++G(d,p)** and **WB97XD/CBSB7** methods. The **WB97XD** is a **DFT** method, which is recommended for modelling *π-stacking* interactions[160].

- The calculation types performed are listed in *5.6*, where the red cross means the calculations did not converged; therefore, only single-point energy calculation was performed for the molecule with fixed geometric parameters. For the *thienyl*-derivatives, our choice was 300 pm bond length for the secondary *H-bond*, and 350 pm for the *selenophenyl* derivatives. These bond lengths are based on lower level calculations and the *Bondi* criteria.[163]

5. RESULTS AND DISCUSSION

5.1 Summary of the synthetic part

5.1.1 The results of the E-selective methods

Several different condensation reactions were tried to find the best method having the highest yield and stereoselectivity for the synthesis of *heteroaromatic cinnamic acid* derivatives. The results of the whole synthetic part are summarized in *Table 7*. As one can observe, the yield and the selectivity of the listed methods leading to *E-cinnamic acid* show very little difference; nevertheless, these techniques differ from each other in the nature of the products, the difficulty of the reaction and the number of the reaction steps. The *Knövenagel-Döbner* condensation was chosen for the synthesis of *E-heterocyclic cinnamic acids*, because this reaction provides the best yield and reproducibility in a one-pot synthesis, and the purity of the resulting *cinnamic acids* was fair enough without recrystallization (>95%), and after a single purification step, it increased to 99–99.5%. The disadvantages lie in the solvent and the catalyst, because pyridine **43** and piperidine **20** are noxious materials. The *Knövenagel-Döbner* method has very good scalability as well. The yields did not drop below 95% on increasing the starting amount of *benzaldehyde* from 50 µl to 50 ml. Therefore, this method was applied for the synthesis of all the targeted materials, and, as it was indicated in the experimental section (*4.3.7*), the yield and the selectivity were very good in all cases. Consequently, it can be stated that the *Knövenagel-Döbner* method is the best choice for the synthesis of *E-cinnamic acids* and *E-3-heteroaryl propenoic acid* derivatives. It is especially true for the *pyridylcinnamic acids*, because they are hardly accessible via other condensation methods.

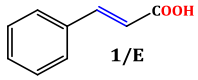
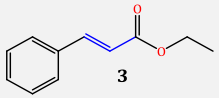
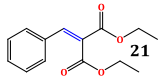
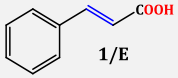
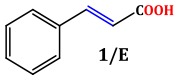
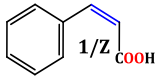
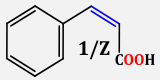
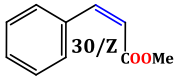
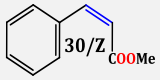
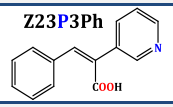
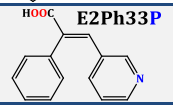
Condensation Method	Reaction time	Product	Yield	*Z:E* ratio	Additional steps to title product	Advantages	Disadvantages
Perkin	2 h		81-92%	0:100	2	Simple setup, non-toxic reagents, acceptable yield.	Side reactions with some heterocyclic aldehydes.
Claisen-Schmidt	4 h		60-65%	0:100	2	Cheap reagents, leads to esters in one step.	Handling molten sodium, high solvent volume needed, low yield.
Knövenagel	14 h		85-87%	0:100	3	The first step leads to malonic acid derivatives.	Complicated setup, long reaction time, 3 additional steps needed.
Knövenagel-Döbner	3 h		95-99%	0:100	1	Simple setup, excellent yield in one step, high purity, cheap reagents.	Noxious solvent and catalyst.
Perkin-Döbner (MW)	30 min		60-87%	0:100	2	Very short reaction time, simple conditions.	Expensive reagents and safe microwave reactor needed.
UV-isomerization	16 h		100%	62:38	0	Simple conditions, acceptable yield for the *Z*-isomer.	The separation of the isomer mixture is difficult.
Lindlar *Z*-selective hydrogenation	2 h		85-95%	95:5	4	High yield and selectivity.	Sensitive for conditions, needs an alkyne derivative as a substrate.
***Z*-selective HWE: Still-Gennari**	3 h		85-95%	84:16	9	High yield and *Z*-selectivity.	Cryogenic conditions needed, the reagent is expensive.
***Z*-selective HWE: Ando**	3h		80-97%	90:10	9	High yield and *Z*-selectivity.	Cryogenic conditions needed, the reagent is hardly accessible.
Modified Perkin condensation	1 h		5%	78:22	5	Simple reaction conditions.	Low yield, the separation is complicated.
HWE: with aromatic phosphonates	2 h		78%	0:100	7	High yield and selectivity.	The reagent is hardly accessible.

Table 7 Summary of the tested reactions for the synthesis of *E* and *Z-cinnamic acids*

5.1.2 Photoisomerization of E-cinnamic acids

The UV-induced isomerization technique is a promising procedure to synthesize *Z-cinnamic acids*. It can be carried out with simple equipment (*4.2.3*) and provides high amount of *Z* isomer within acceptable time scale (12 h for **1/E**). It leads to equilibrium, and can be described as a pseudo first order reaction (*Figure 46*). The equilibrium constant is K = 1.63 ± 0.01 and the rate coefficients are: k_f = 2.37×10^{-5} ± 9.5×10^{-7} s^{-1}, k_b = 1.46 × 10^{-5} s^{-1} ± 1.46 × 10^{-7} s^{-1} for **1/E**.

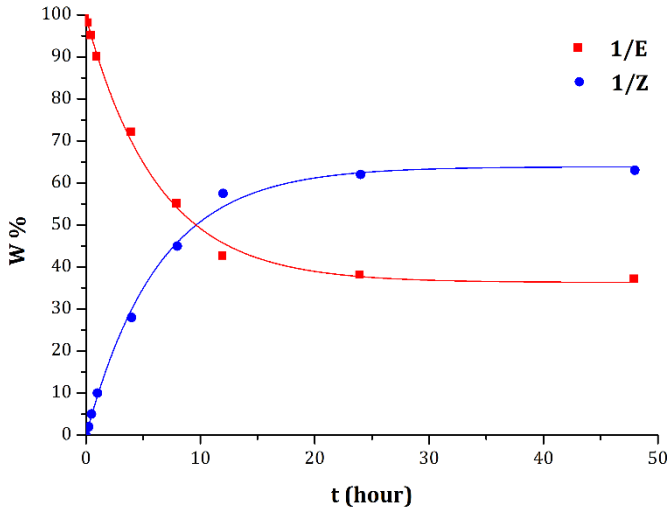


Figure 46 Determining the equlibrium and the rate constants for the isomerization of **1/E** from the kinetic curve

As it was seen in the experimental part (*4.4.1*), the distribution of the isomers in the equilibrium, varies from 29 to 87% (determined by ^{1}H NMR after 24 h irradiation) for the *heteroaryl propenoic acids*; only the **E33P**, **E34P** and **E2Ph33P** derivatives show low isomerization activity because of the low solubility of the compounds in EtOAc. These materials are only soluble in dipolar aprotic solvents like DMSO, which has an absorption band around 254 nm; therefore, the procedure cannot be carried out in DMSO.

The semi discontinuous UV-reactor is applied for the accumulation of **E-Z32F**, **E-Z32T** and **E-Z32P** isomer mixtures with the method described in (*4.2.3*). Several different separation methods were performed on the isomer mixtures to isolate the pure *Z isomers* from the more easily accessible *E* isomer.

One can think about many different techniques to separate the stereoisomers like gradual acidification using the different pKa values of the isomers, or the sublimation

method utilizing the bp. difference among the isomers. The former method did not work with the *heteroarlycinnamic acid* isomer mixtures, while the more promising latter one would have needed special equipment. Simple flash column chromatography failed to separate the isomers as well, even the 50m long capillary GC column could barely separate the *E-Z-methyl ester* derivatives – there was some overlap among the peaks –, nevertheless, the two isomers could be clearly identified, and the determined isomer ratio was the same as was found by ^{1}H NMR spectroscopy (with ± 1% deviation).

To sum up, the isomerization method, combined with isomer separation via sublimation/distillation, can be a relevant method to yield *Z-cinnamic acid* derivatives.

5.1.3 Summary of the Z-selective methods

The various *Z*-selective hydrogenation methods showed excellent yields (85–95%) and *Z*-selectivities (90–96%); however, the reaction used to synthesize the starting materials, like *3-phenylpropiolic acid* **26**, works only for the *phenyl* derivatives of *cinnamic acid*. The *furyl-*, *thienyl-*, and *selenienyl* derivatives yield perbrominated molecules after bromination, even if the reaction was carried out at low temperature. The *pyridyl-derivatives* are insoluble in most of the solvents applied for halogenation, thus, bromination does not take place. The *Heck-Mizoroki* coupling may be an alternative to synthesize the starting alkynes; unfortunately, the reagents are very expensive for this reaction.

The best method to produce *Z-3-heteroaryl propenoic acids* was the *Z*-selective HWE reaction, which is different from the other methods leading to *Z* isomers. Due to the steric hindrance (originating from the large side groups), and the stabilized *carbanion*, the reaction produces *Z-cinnamic acids* with high yield (>90%) and stereoselectivity (>90%), as it is shown in *4.5.5*. The route to the reagents is long (1-2 week), but if one searches for appropriate methods, it is possible to access the reagents at acceptable cost. The method requires special conditions, but if the reaction is well controlled, the *Z*-selectivity can be as high as 97%, and the method works well with each *heterocyclic carboxaldehyde*; therefore, the HWE reaction, applying the *Ando* reagent is found to be a general method for the preparation of *Z-3-heteroaryl cinnamic acids*.

5.1.4 Summary of the synthetic methods for α-substituted cinnamic acid derivatives

As it was mentioned in the experimental part, the *thienyl-substituted* variants were available through the modified Perkin condensation with acceptable yield (34-67%). The stereoismers can be separated via column chromatography and selective precipitation methods.

The modified Perkin reaction was also tested for the preparation of pyridyl-substituted derivatives, and *Z-2-(3-pyridyl)-3-phenylpropenoic acid* could be synthesized successfully, although with unacceptably low (5%) yield. Therefore, other methods were tried and the HWE reaction seemed to be a choice in this case as well. The yield was 78% for *E-Z-2-phenyl-3-(3-pyridyl)propenoic acid*, and the isomer distribution was ?. The general approach to the HWE reagent (both *pyridyl-* and *phenyl phosphonoacetates*) are similar to the already mentioned *Ando* and *Still-Gennari* reagents, and the route to the compound was also long; nevertheless, the final condensation reaction performed well and the processing of the reaction mixture was uncomplicated, thus, this method is more suitable for the preparation of the *pyridyl derivatives* than the modified *Perkin* condensation. To sum up, the HWE method can be applied for the synthesis of every possible *E-Z-2-(x-pyridyl)-3-(x-pyridyl)propenoic acid*.

5.2 Determining the different hydrogen bonds with infrared spectroscopy

IR spectroscopy was applied to identify the strong cyclic double C=O…H–O *hydrogen bonds* between the *carboxylic* groups. The very typical, wide O–H stretching vibration shows up in the 2800–3500 cm^{-1} region when the O–H…O *hydrogen bonds* are present, and forms a sharp peak around 3600 cm^{-1} when the O–H group is not involved in hydrogen bonding. We know from our previous studies that each *cinnamic acid* derivative forms *hydrogen-bonded* cyclic dimers, just like any other *carboxylic acid*, in the solid state and also in the liquid phase when aprotic solvent is used. The weaker *hydrogen bonds* can be identified from the displacement of the (aromatic)C–H…X and/or (olfenic)C–H… X vibrations; however, these changes may also belong to *π-stacking* interactions.

The solid state infrared spectra of the studied materials are shown from *Figure 47* to *Figure 50*.

The spectra of the *furyl*, *thienyl* and *selenophenyl* derivatives (*Figure 47*) look very similar: the pristine acids (both *E* and *Z* isomers) form *hydrogen-bonded* dimers (O–H…O=C) indicated by the wide O–H stretching vibration appearing around 2500–3000 cm^{-1}, and the same vibration disappears if the acidic protons are no longer present (in the *sodium salt* of the materials). All studied materials have common vibrations regardless of the form (*acid* or *sodium salt*) like the strong C=O stretching at

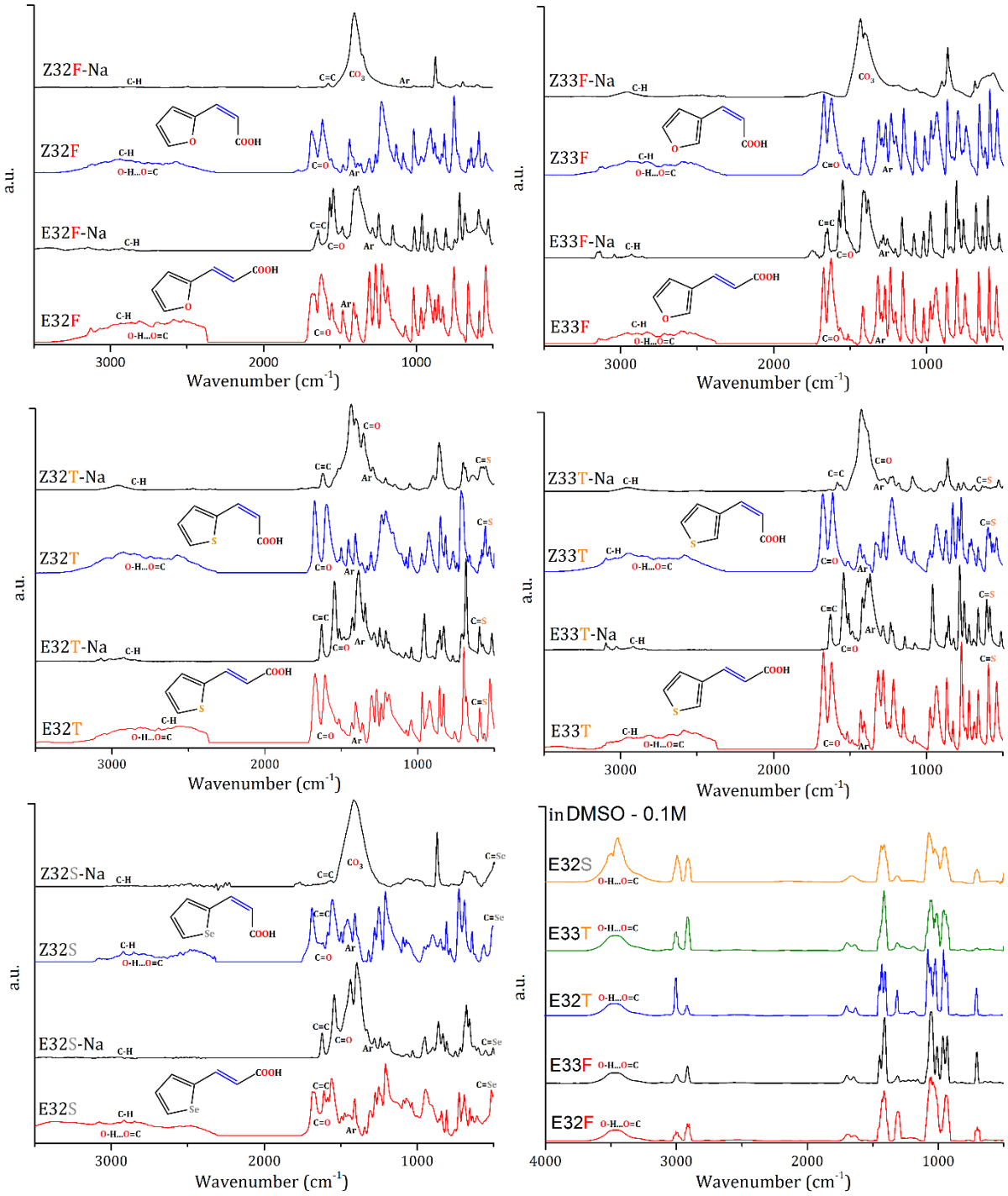


Figure 47 The solid and liquid state IR spectra of the *furyl-, thienyl- and selenophenyl*-derivatives

1600 cm^{-1}, the olefinic C=C at 1700 cm^{-1}, the aromatic skeleton vibrations under 1500 cm^{-1} or the C–H stretching around 3000 cm^{-1} as it indicated in the spectra. Some identical vibrations can be assigned as well, like the different bending of the aromatic rings under 800 cm^{-1}, where the *heteroatom–carbon* atom deformations provide with assignable peaks [C–S, C–Se, Pyr(*pyridine* ring bending)].The liquid-state spectra of the *E* isomers, in DMSO are registered as well to check the presence of the double *hydrogen-bonds* in solution. As it is seen, the wide absorption band, indicative of *hydrogen bonding* interaction, appears around 3500 cm^{-1}. Sharp O–H vibrations were not found, which means the fundamental unit is the dimer in aprotic solvents, thus, *hydrogen bonding* interactions can be studied in aprotic dipolar media like DMSO. However, the absorption of the solvent is very intense in a wide frequency range, overlapping with the aromatic region making the identification of vibration band displacement at least very difficult if not impossible. A compromise, working in many cases, is the use of *chloroform*, which has a spectrum significantly poorer in vibrations in the mid-IR range, and can dissolve many but not all of our compounds in appreciable amounts. Unfortunately, some of the molecules studied are nearly insoluble in this solvent (*pyridyl*- and *α-substituted pyridyl* derivatives). In these cases, DMSO was chosen as the solvent, and NMR spectroscopy was applied for studying *hydrogen bonding* interactions instead of IR spectroscopy.

The band indicative of the *hydrogen bonds* between the *carboxylic* groups disappeared from the spectra of the pristine *pyridylpropenoic acids* (*Figure 48*) (both *E* and *Z* isomers), and a new vibration showed up at 2500 cm^{-1}, which can be assigned to quaternary N^{+}–H stretching. This means the *pyridyl*-derivatives exist as *zwitterions*, which not so surprising, because the *pyridyl*-ring has basic character and a proton is easily transferred from the *carboxylic* group of another *pyridylpropenoic acid*. Only the **E32P** variant shows some dimerization. In the *hydrochloric salt*, the *carboxylate* group is forced to keep its proton, thus, dimerization occurs again, as it is seen in *Figure 48*.

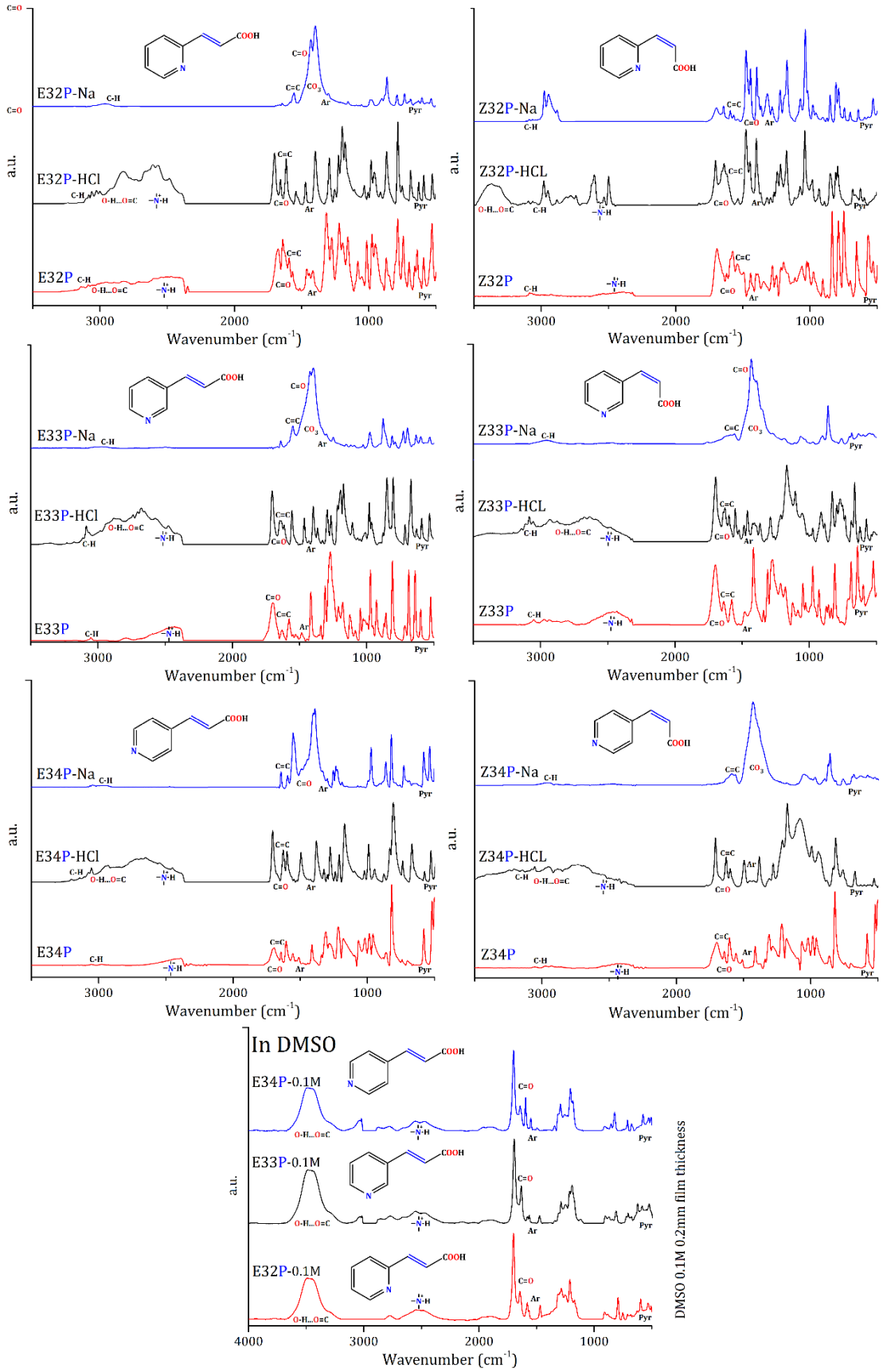


Figure 48 The solid- and the liquid-state IR spectra of the *pyridylcinnamic acids*

The *α-substituted thienylcinnamic acid* derivatives show very similar results to those seen for the unsubstituted cinnamic acid analogs containing five-membered heterocycles. The typical wide O–H…O=C vibration, the sign of *hydrogen-bonded* dimers, appears in all cases, and the shape of the absorption shows consistency with the geometry of the proper isomer (*Figure 49*). The changes in the shape of the band are rooted in the presence or lack of steric hindrance between the molecules. If the aromatic rings are on the same side of the molecule, they do not interfere with the other parts of the molecule during the formation of the dimer; however, if they are on different sides (*Z* isomers), the intermolecular interactions become predominant, leading to a more regulated crystal structure, yielding narrow O–H..O=C vibration. The spectra of the compounds were registered in solution as well. The solvent was *chloroform* and the concentrations varied in the 10^{-2} – 10^{-4} range. As one can see in *Figure 49*, the aromatic vibrations and the C=S–C vibrations of the *thienyl* group are shifted with 5–10 cm^{-1} to the lower energies with dilution, which means that aromatic and/or olefinic C–H…S interactions prevail in concentrated solutions and in the solid state.

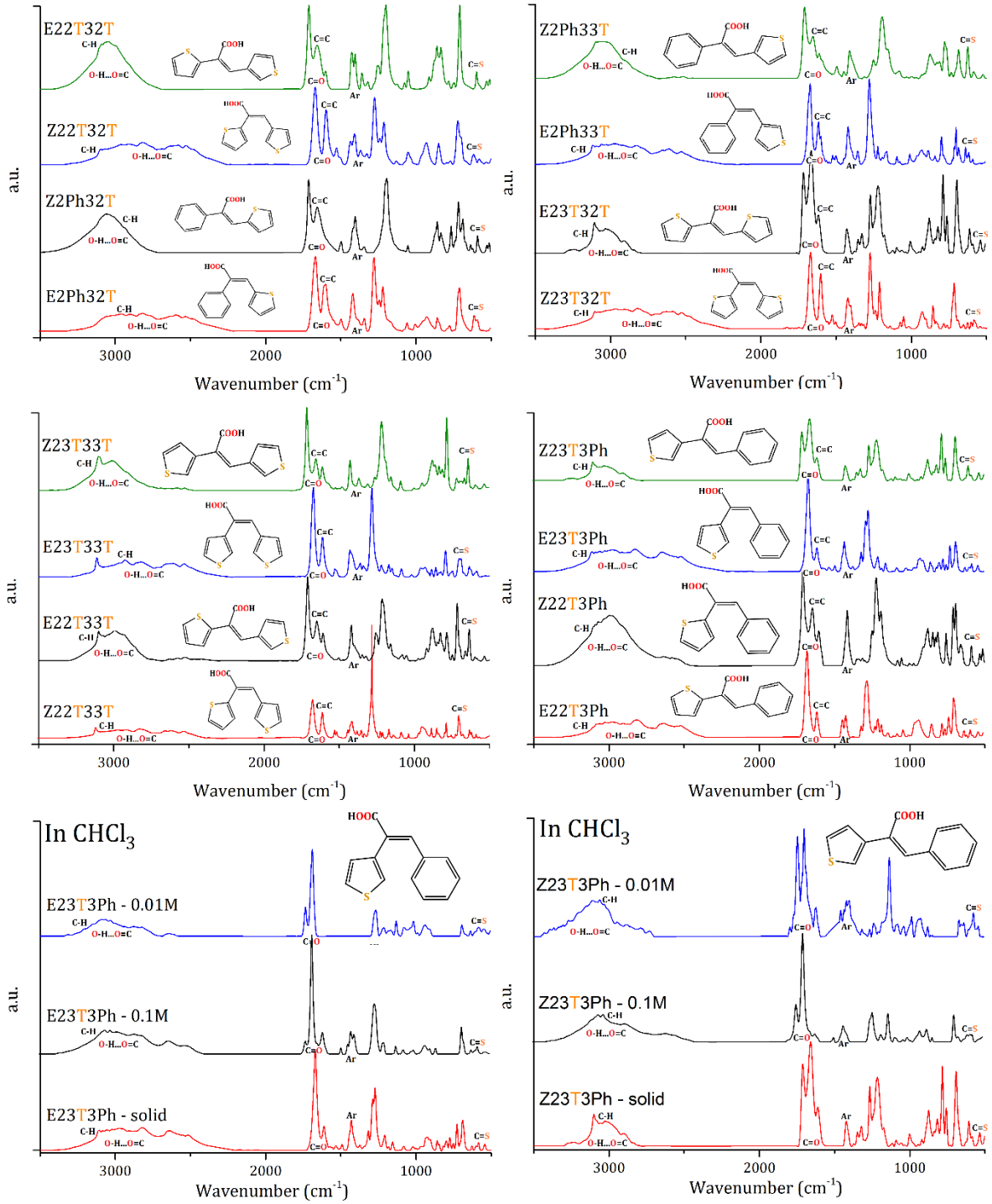


Figure 49 The solid- and the liquid-state IR spectra of the *α-substituted thienylcinnamic acids*

The family of the *α-substituted pyridylcinnamic acid* molecules shows full consistency with the *pyridylcinnamic acids*. The IR spectra of the studied derivatives are shown in *Figure 50*. The practically exclusive structure in the solid state is the *zwitterion*, just as it was seen for the (unsubstituted) *pyridylcinnamic acids*. However, for the **E2Ph33P** isomer dissolved in DMSO, dimerization is the exclusive process (the vibration

of the *quaternary nitrogen* vanished from the spectra); therefore, aggregation via the weaker C–H…N interactions becomes possible.

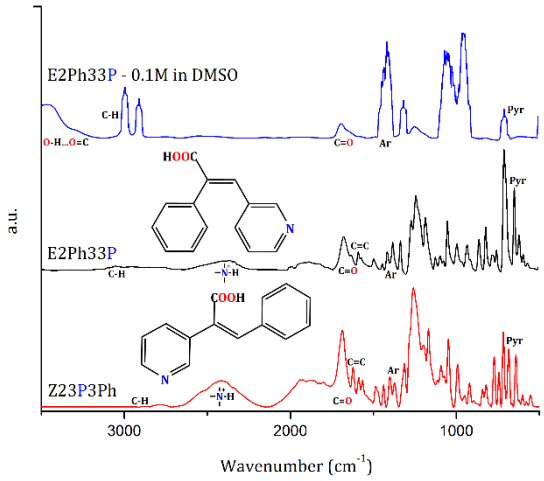


Figure 50 The solid- and liquid-state IR spectra of the *α-substituted pyridylcinnamic acids*

To sum up, it has been found that the fundamental unit is the double *hydrogen-bonded* cyclic dimer in the solid state for the *five-membered* derivatives and the *zwitterion* for the *pyridyl* derivatives; however, they can form dimers in DMSO; therefore, the further investigations the interactions between the dimers are feasible. Based on the experimental findings, the *hydrogen-bonded* dimers are used as fundamental units in the quantum chemical modelling as well.

5.3 ^{1}H–^{1}H correlation (COSY) and nuclear Overhauser effect (NOESY) NMR spectra of heteroaryl cinnamic acids

As was seen in the IR spectra of the studied materials, the fundamental unit is the *hydrogen-bonded* dimers kept together via strong O–H…O=C cyclic *hydrogen bonds* in most cases. The *pyridyl* derivatives only act differently in the solid state; nevertheless, they also form the cyclic dimers in dipolar aprotic solvent like DMSO.

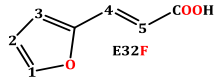
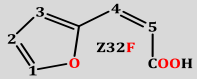
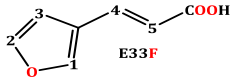
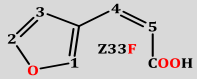
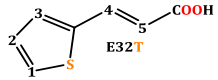
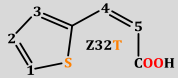
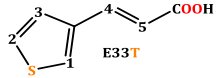
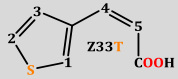
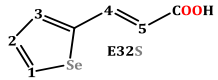
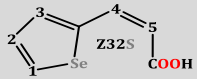
Compound	H1	H2	H3	H4	H5	-COO-H
	8.19, d J = 1.2 Hz	7.00, dd J_A= 1.5 Hz, J_B = 2.8 Hz	7.29-7.30, d J = 2.7 Hz	7.79-7.82, d J = 15.5 Hz	6.57-6.61, d J = 15.5 Hz	12.77
	7.76, d J = 1.1 Hz	6.60, dd J_A = 1.3 Hz, J_B = 2.5 Hz	7.50-7.51, d J = 2.7 Hz	6.79-6.81, d J = 12.8 Hz	5.73-5.75, d J = 12.8 Hz	12.41
	8.02, s (br)	7.67, s (br)	6.87, s (br)	7.47-7.50, d J = 15.8 Hz	6.20-6.23 J = 15.8 Hz	12.35
	7.67, s (br)	8.22, s (br)	7.02, s (br)	6.77-6.79, d J = 12.8 Hz	5.74-5.77 J = 12.8 Hz	12.31
	6.86-6.87, d J = 4.2 Hz	6.30-6.31, dd J_A = 2.5 Hz, J_B = 3.9 Hz	6.67, d J = 2.7 Hz	6.89-6.92, d J = 15.8 Hz	5.33-5.36, d J = 15.8 Hz	11.55
	7.73, d J = 3.2 Hz	7.09, dd J_A = 3.1 Hz, J_B = 3.4 Hz	7.53, d J = 3.6 Hz	7.19-7.22, d J = 12.2 Hz	5.69-5.72, d J = 12.2 Hz	12.37
	7.52-7.53, dd J_A = 4.8 Hz, J_B = 2.6 Hz	7.86, d J = 2.3 Hz	7.47-7.48, d J = 4.8 Hz	7.58-7.62, d J = 16.1 Hz	6.34-6.37, d J = 16.1 Hz	12.27
	7.55, s (br)	8.10, s (br)	7.51, s (br)	6.88-6.91, d J = 12.5 Hz	5.77-5.80, d J = 12.5 Hz	12.35
	8.28, d J = 4.5 Hz	7.31, s (br)	7.63, s (br)	7.72-7.75, d J = 15.2 Hz	6.07-6.10, d J = 15.2 Hz	12.39
	8.25, d J = 5.1 Hz	7.59-7.60, dd J_A = 5.4 Hz, J_B = 3.3 Hz	7.37-7.38, d J = 3.5 Hz	7.28-7.31, d J = 11.9 Hz	5.68-5.70, d J = 11.9 Hz	12.20

Table 8 ^{1}H NMR assignements for the *five-membered heteroaryl cinnamic acid* derivatives in DMSO-d_6

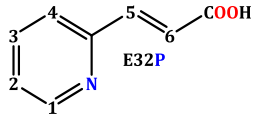
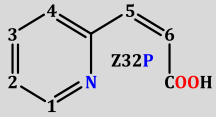
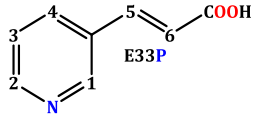
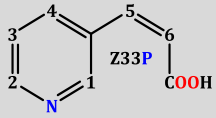
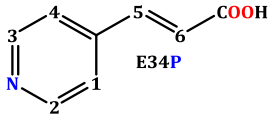
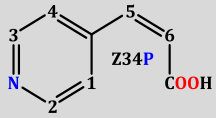
Compound	H1	H2	H3	H4	H5	H6	-COO-H
 E32P	8.59-8.60, d J = 4.2 Hz	7.40-7.43, ddd J_A = 4.6 Hz, J_B = 6.4 Hz,(J_C = 0.9 Hz)	7.87-7.89 ,dt J_A = 7.7 Hz, J_B = 6.2 Hz	7.68-7.69, d J = 7.6 Hz	7.56-7.59, d J = 15.9 Hz	6.83-6.86, d J = 15.9 Hz	12.43
 Z32P	8.71-8.73, d J = 4.2 Hz	7.59-7.61, dd J_A = 4.7 Hz, J_B = 5.7 Hz	8.12-8.13, dt J_A = 7.5 Hz, J_B = 5.8 Hz	7.76-7.78, d J = 7.6 Hz	7.03-7.06, d J = 13.1 Hz	6.14-6.17, d J = 13.1 Hz	12.22
 E33P	8.82, s (br)	8.55-8.56, d J = 4.2 Hz	7.40-7.42, dd J_A = 7.6 Hz, J_B = 4.5 Hz	8.11-8.13, d J = 7.8 Hz	7.59-7.62, d J = 16.1 Hz	6.64-6.67, d J = 16.1 Hz	12.54
 Z33P	8.67, s (br)	8.49-8.48, d J = 3.5 Hz	7.37-7.38, d (br), J = 6.8 Hz	8.00-8.02, d J = 7.4 Hz	6.95-6.98, d J = 12.5 Hz	6.06-6.09, d J = 12.5 Hz	12.53
 E34P	7.49-7.52, d J = 3.9 Hz	7.60-7.61, d J = 3.8 Hz	7.60-7.61, d J = 3.8 Hz	7.49-7.52, d J = 3.9 Hz	7.49-7.52, d J = 16.1 Hz	6.70-6.74, d J = 16.1 Hz	12.36
 Z34P	8.41-8.42, d J = 4.1 Hz	7.20-7.21, d J = 4.2 Hz	7.20-7.21, d J = 4.2 Hz	8.41-8.42, d J = 4.1 Hz	6.86-6.89, d J = 12.2 Hz	6.15-6.17, d J = 12.2 Hz	12.48

Table 9 ^{1}H NMR assignements of the *pyridylcinnamic acids* in DMSO-d_6

The complete proton assignments of the studied molecules are shown in *Table 8*, *Table 9* and *Table 10*. As it can be seen, the acidic protons of the *carboxyl* groups are present in each spectrum, even for the *pyridylcinnamic acids*, where this peak is detectable for **E32P**, **Z32P**, **E34P** and **Z34P** and is sharp for **E33P**, **Z33P, E2Ph33P** and **Z23P3Ph**. This is another evidence for the presence of the cyclic dimers in the solution. Moreover, the peaks of the aromatic *protons* are unusually broadened (deconvolution was needed sometimes to resolve the broad peaks) compared to the phenyl derivatives, and ddd (doublets of double doublets) couplings, which would be a sign of complex interactions, were very rarely found (**E32P**). However, the signal broadening and the dislocation of the peaks did not provide enough evidence for the presence of weak *hydrogen bonds*; therefore, $^{1}H-^{1}H$ *correlation spectroscopy* (COSY) was chosen to clearly identify interactions like these. The COSY method, reveal the couplings between the neighboring protons as strong, off-diagonal peaks. Coupling with the second and third neighbor can be detected as well; however, these peaks are very small. If weak *hydrogen bonds* are present among the molecules, couplings with the second or the third neighbor (in the *heteroaromatic* rings) become stronger. To avoid confusion between couplings originating from covalently-bonded and *hydrogen-bonded* systems, coupling between the first olefinic *proton* and the closest aromatic *proton* (H3–H4 for the *five-membered* derivatives and H4–H5 for the *six-membered* variants) was chosen as internal reference. The signal level of these coupling was the threshold of the measurements. Every coupling having lower intensities than the reference peak was removed from the spectra (including the reference peak as well) to simplify the understanding of the interactions.

Figure 51 shows the results for the *furylcinnamic acid* derivatives. Through the combinations of the couplings, only one possible *hydrogen-bonded* structure can be identified for *3-(2-furyl)cinnamic acids*, where the coupling among H1–H3 belongs to the (aromatic)C–H…O–(aromatic)C weak *hydrogen bond*. For the *3-(3-furyl)* derivatives, one more interaction can be assigned via the H1–H2 and the H1–H3 couplings.

The $^{1}H-^{1}H$ COSY spectra of the *thienyl* variants are shown in *Figure 52*. The couplings assigned are as follows: H1–H3 for *3-(2-thienyl)*, H1–H2 and H1–H3 for the *3-(3-thienyl)* derivatives. Although the (aromatic)C–H…S–(aromatic)C interactions are significantly weaker than the (aromatic)C–H…O–(aromatic)C *hydrogen bonds*, it is still clearly detectable.

The weakest interaction was found for the *3-(2-selenophenyl)* compounds (*Figure 53*), where the signal level of the H1–H3 coupling is hardly reaches the selected threshold, which means the (aromatic)C–H…Se–(aromatic)C *hydrogen bond* is the weakest in the *heteroarylcinnamic acid* families studied.

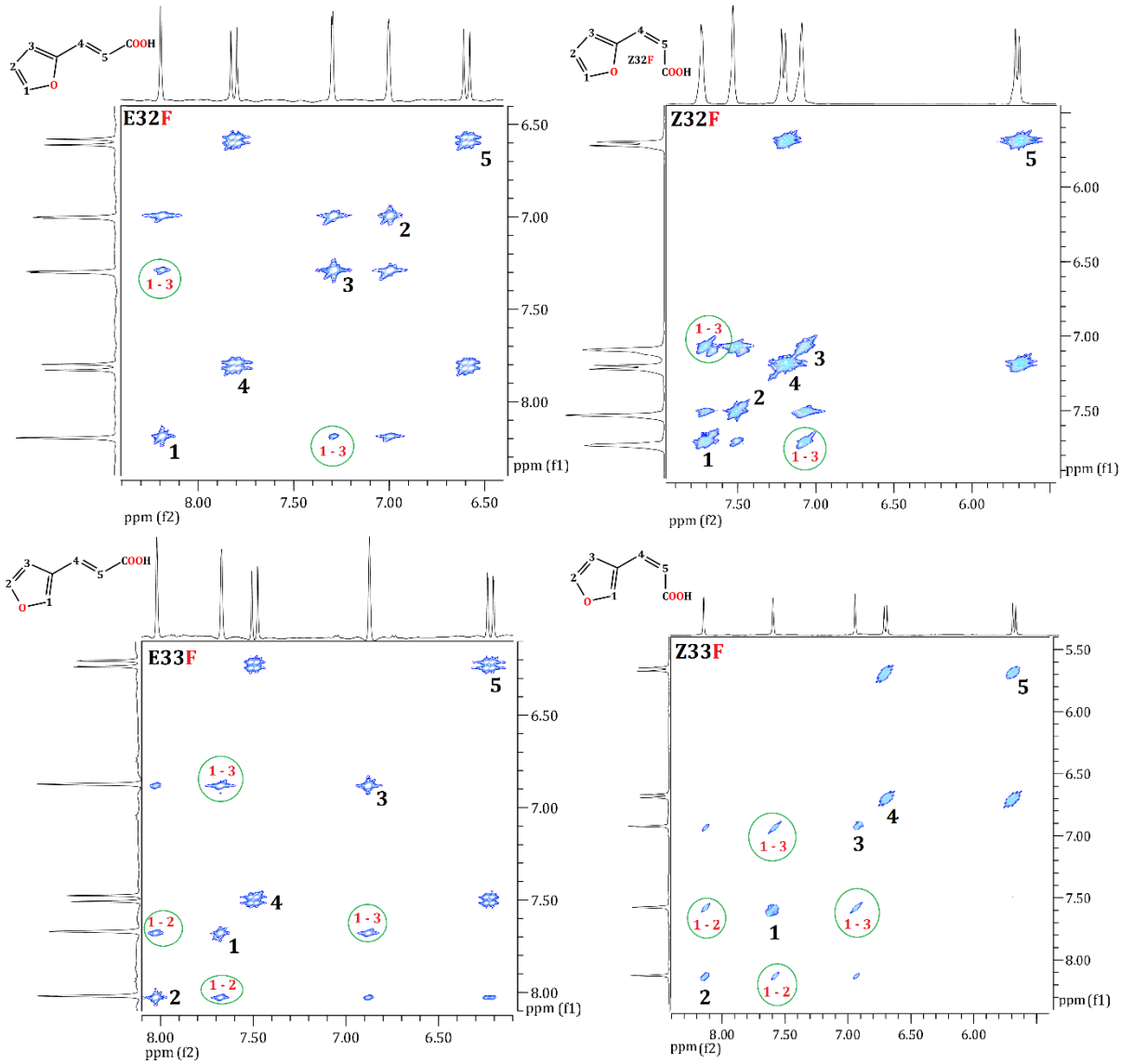


Figure 51 The $^{1}H–^{1}H$ COSY NMR spectra of the *furylcinnamic acid* derivatives

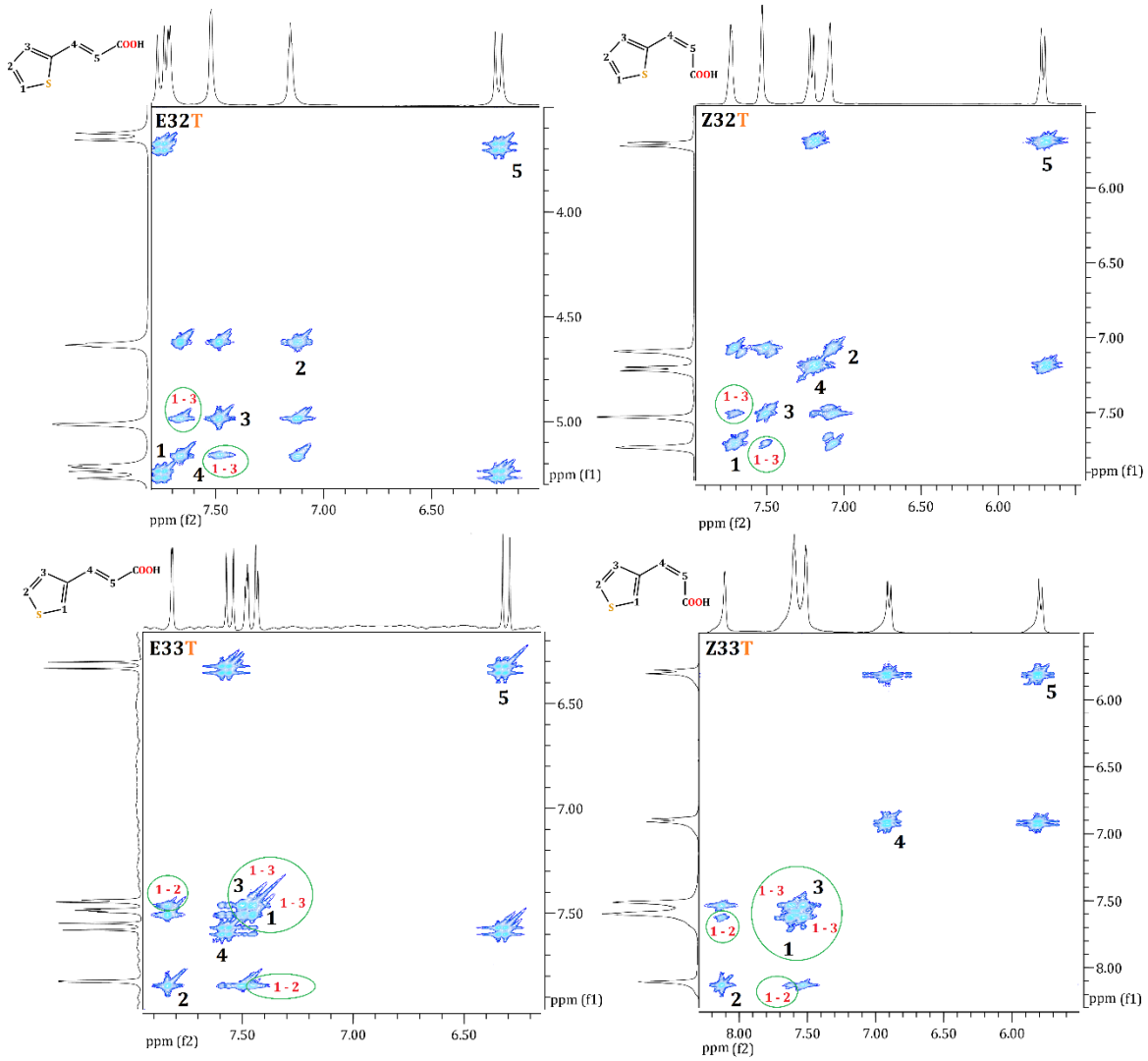


Figure 52 The $^{1}H-^{1}H$ COSY NMR spectra of the *thienylcinnamic acid* derivatives

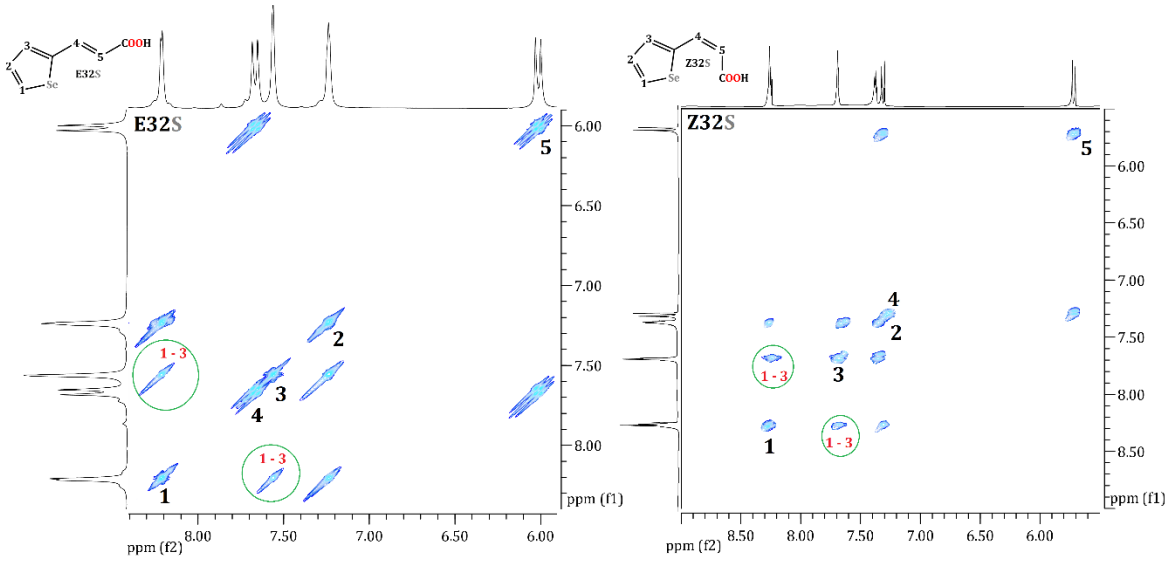


Figure 53 The $^{1}H-^{1}H$ COSY NMR spectra of the *selenophenylcinnamic acid* derivatives

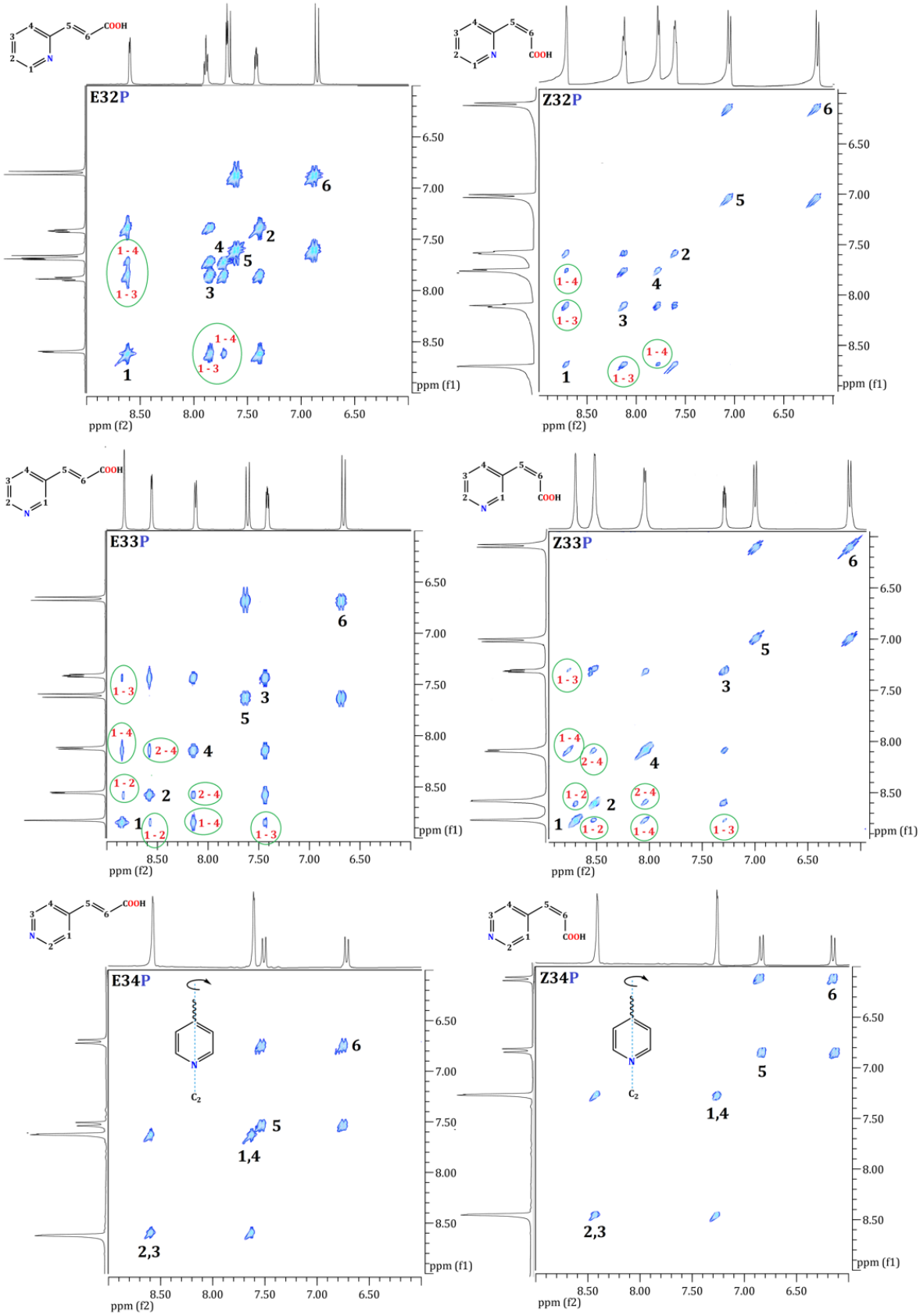


Figure 54 The $^{1}H-^{1}H$ COSY NMR spectra of the *pyridylcinnamic acid* derivatives

The *pyridyl* derivatives show much more (aromatic)C–H…N–(aromatic)C interactions (*Figure 54*) due to the bigger ring size. H1–H3 and H1–H4 couplings are found with intense signals for the *3-(2-pyridyl)* variants and H1–H2, H1–H3, H1–H4 and H2–H4 couplings with the same parameters for the *3-(3-pyridyl)* derivatives. Unfortunately, the *pyridine* ring of the *3-(4-pyridyl)* isomers have C_2 symmetry (*Figure 54*), which means the H1, H4 and H2, H3 protons are magnetically equivalent. Thus, they show only two signals, and these *protons* are in strong *spin-spin* couplings, which completely cover any other interactions. However all signals are quite broad, which is a sign of *hydrogen bonds*, and compared to other *pyridyl* derivatives, these molecules are the less hindered variants; therefore, the possibility to form cyclic tetramers connected by (aromatic)C–H…N–(aromatic)C weak *hydrogen bonds* are very feasible.

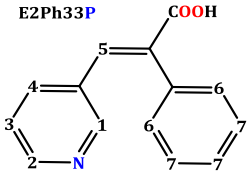
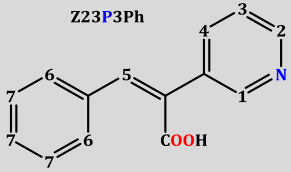
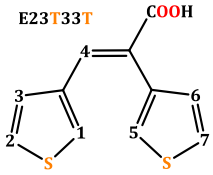
Compound	H1	H2	H3	H4	H5	H6	H7	-COO-H
	7.82, s	7.86-7.87, d J = 3.3 Hz	6.68-6.70 ,dd J_A = 7.6 Hz, J_B=3.8 Hz	6.86-6.87, d J = 7.8 Hz	7.08, s	6.77, m	7.03, m	12.98
	8.33, s	8.54-8.55, d J = 3.8. Hz	8.41-7.43, dd J_A = 7.5 Hz, J_B = 3.9. Hz	7.76-7.78, d J = 7.6 Hz	7.90, s	7.06-7.07, m	7.23-7.26, m	12.96
	7.40, s	7.26, br	7.00-7.01, d J_A = 4.0 Hz,	7.95, s	7.14, s	6.60-6.61, d J = 4.0 Hz	7.24, s	12.29

Table 10 ^{1}H NMR assignements of the studied *α-substituted cinnamic acids* in DMSO-d_6

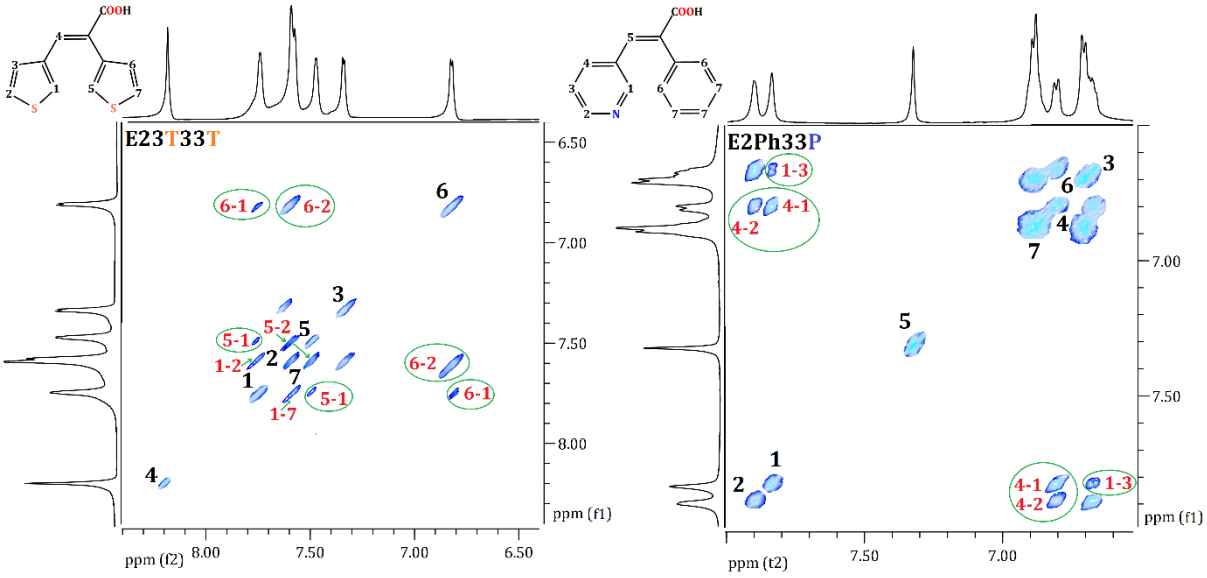


Figure 55 The ^{1}H-^{1}H COSY NMR spectra of two *α-substituted cinnamic acid*

The previously discussed interactions were found for the *α-substituted* derivatives as well; however, it was not possible to identify each feasible interaction because of the overlapping peaks. For the **E23T33T** derivative, the couplings among H1–H2, H1–H5, H1–H6, H1–H7 and H2-H6 protons indicated *hydrogen bonding* interactions. The other example is the **E2Ph33P** derivative, where the H1–H3, H1–H4 and H2–H4 couplings, indicating *hydrogen bonding* interactions, could be identified clearly, and there are ambiguous hints for other interactions as well, e.g. the *protons* of the *phenyl* ring may be involved in the *hydrogen-bonded* system, but due to the overlapping signals we cannot be sure about this interaction.

Utilizing the ^{1}H–^{1}H *nuclear Overhauser effect spectroscopy* (NOESY) it was possible to explore the conformational space and properties of the molecules dissolved in various solvent types (DMSO – dipolar, aprotic, CD_3OD – polar, protic, $CDCl_3$ – close to being completely apolar) using long-range couplings. For *cinnamic acids*, the coupling between the olefinic and aromatic proton(s) was used to deduce the conformational behavior of the aromatic rings. The results for the first studied molecule, **Z23P3Ph** that is, are shown in *Figure 56*. The spectra were recorded in relatively diluted (0.05 M) solution (unlike the COSY experiments) to minimize the possibility of intermolecular interactions. As it is seen, the only clear coupling is between H5–H6. This means that the *pyridyl* ring can rotate freely and/or it is far from the completely flat structure; while the

rotation of the phenyl group is hindered or the structure in equilibrium is close to being flat.

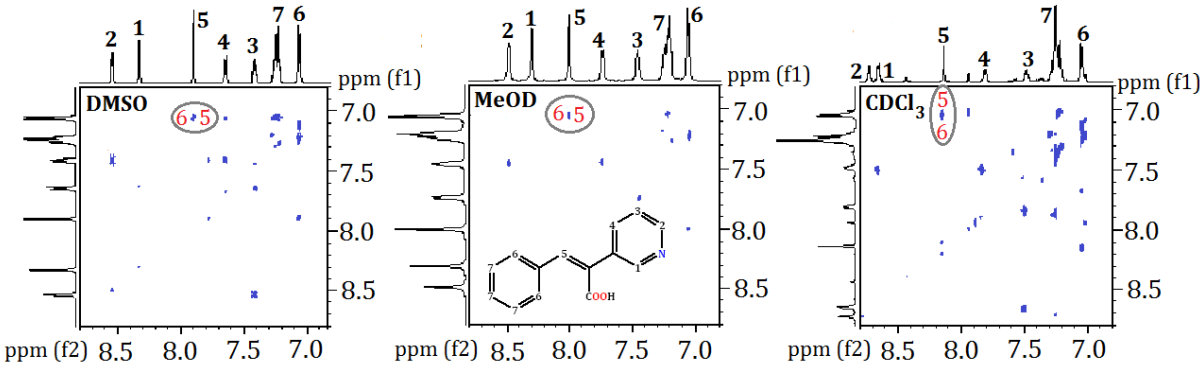


Figure 56 ^{1}H–^{1}H NOESY spectra of **Z23P3Ph** in widely different solvents at 0.05 M concentration

Another example is the ^{1}H–^{1}H NOESY spectra of **E32P** and **E34P** (*Figure 57*). Surprisingly, no cross peaks can be detected in the spectra, which means that the aromatic ring is free to rotate between the several possible minima.

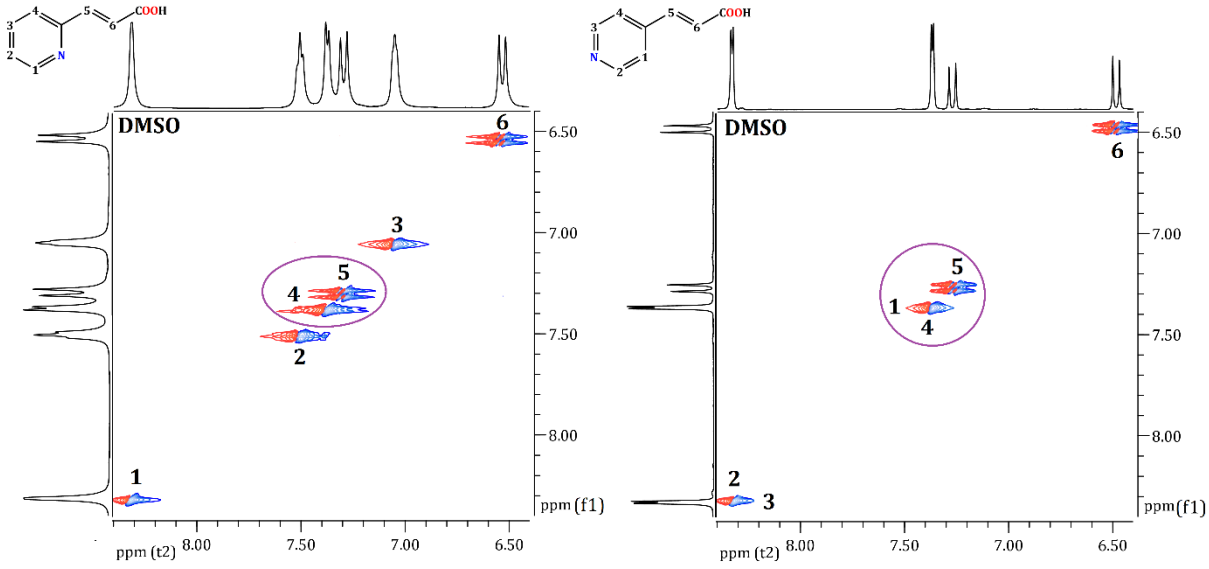


Figure 57 The ^{1}H–^{1}H NOESY spectra of **E32P** and **E34P** in DMSO at 0.05 M concentration

To sum up, it can be stated that many intermolecular *hydrogen bonding* interactions are possible among each *heteroarylcinnamic* and *α-substituted heteroarylcinnamic acids*. IR and/or NMR spectroscopies proved to suitable tools for obtaining qualitative, and the intensities of the crosspeaks in the ^{1}H–^{1}H COSY spectra, semiquantitative information concerning these close contacts. It was found that the

strength of the (aromatic)C–H…X–(aromatic)C is decreases in the C-H…N > C-H…O > C-H…S > C-H…Se order. However, for quantitative (energy and geometric) data, quantum chemical calculations are needed (see later: *5.6.2*). From the results of the NOESY spectra, it has been established that the rotation is nearly free for each derivative studied; however, for the complete description of the conformational properties, quantum chemical calculations are needed once again (see later: *5.6.1*).

5.4 Identifying the π-stacking interactions with powder X-ray diffractometry

The *π-stacking* interactions may be present in the liquid and the solid state as well; however, they become only predominant in the crystalline state for small molecules like *cinnamic acid* derivatives. The appearance of the studied materials are quite uniform, the *five-membered* variants form monoclinic crystals and the *pyridyl* derivatives precipitate from the reaction mixture as rhombohedral crystals due to the *zwitterionic* structure, mentioned earlier (*5.2*).

As it was seen, *cinnamic acids* have a layered structure[95] due to the combination of different *hydrogen bonds* as well as *π-stacking* interactions. To identify the *π-stacking* interactions powder *x-ray diffractometry* (XRD, *4.2.1*) was applied. In the diffractograms, we looking for the signs of the layered structure (interlayer distances in the 350–500 pm range) belonging to *π–π bonding* of the "sandwich" type among the *heteroaromatic* rings. (*Figure 58* and *Figure 59*) display the results for the five-membered variants. It is seen that the *furyl* derivatives have intense reflections in the 15–20 2Θ° region, thus, their structures are very similar, and assuming that these are first-order reflections, the calculated interlayer distances are within the range given above (*Table 11*).

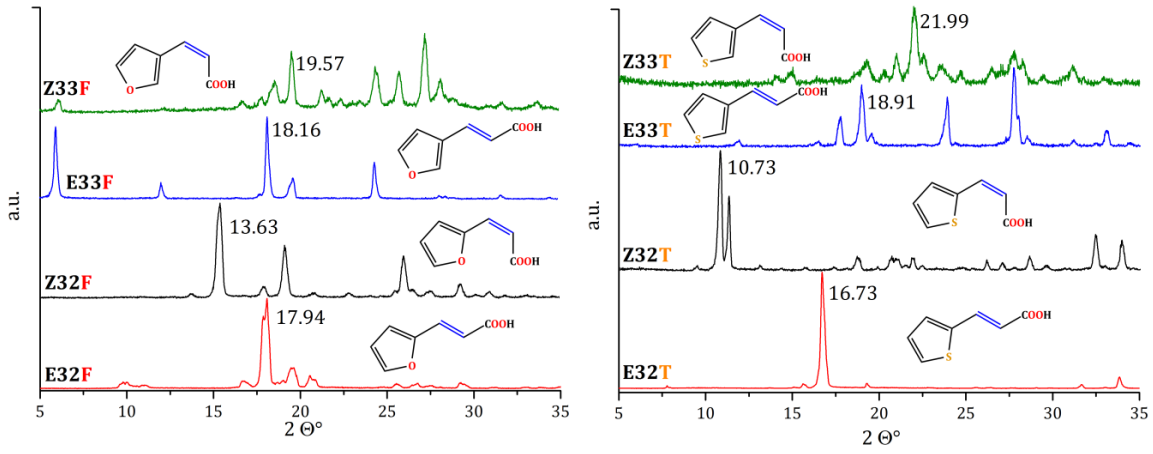


Figure 58 The powder X-ray diffractograms of the *furyl-*, and *thienylcinnamic acids*

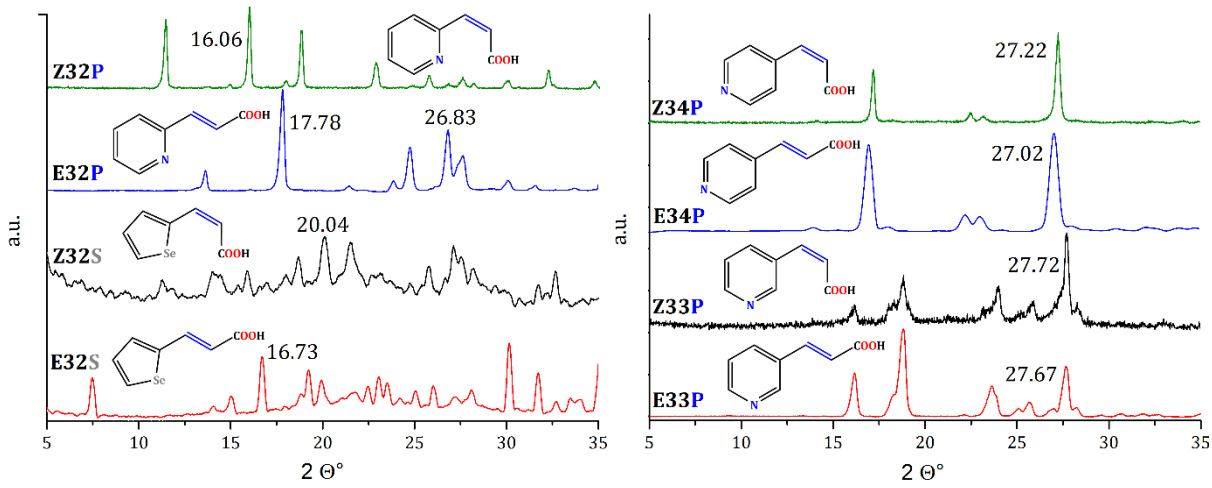


Figure 59 The powder X-ray diffractograms of the *selenophenyl-*, and *pyridylcinnamic acids*

The major interaction for the *pyridyl* derivatives in the solid state is the ionic bond among the *zwitterions*; therefore, these molecules do not have layered structure.

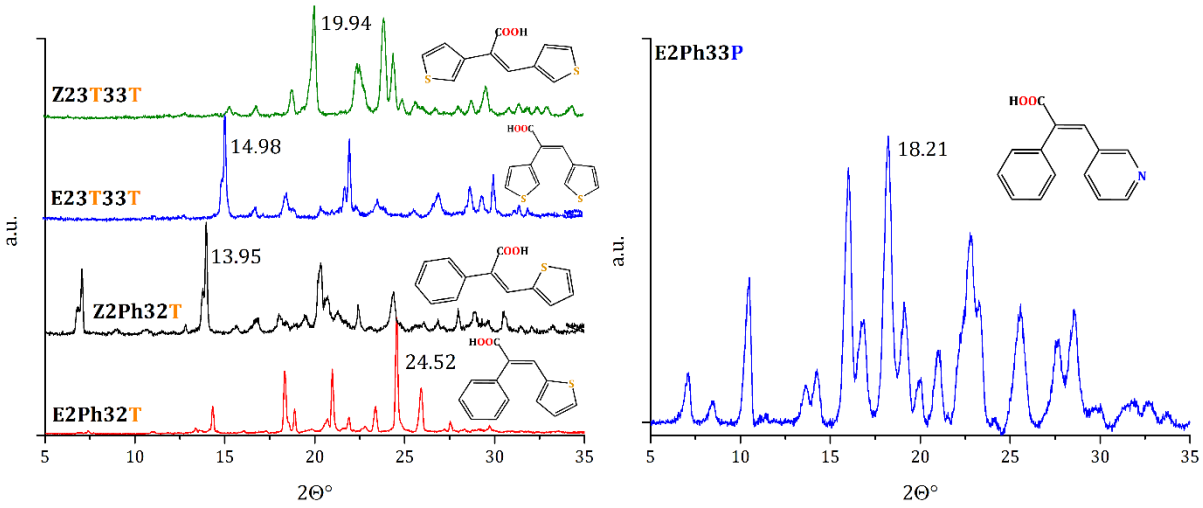


Figure 60 Powder X-ray diffractograms of some *α-substituted heteroaryl cinnamic acid* derivatives

The *α-substituted* derivatives have highly complicated X-ray diffractograms (*Figure 60*), and single crystal data are not available either; therefore, we decided not use the diffractograms for estimating the π−π distances in the assumed but not proven π−π close contacts.

For calculating the interlayer distances, the *Bragg* equation (*Equation 1*) was applied, and selected peak was treated as a first-order reflection (n=1) (unfortunately, the determination if it is a first-order reflection indeed, is difficult from powder X-ray diffractograms; therefore these results are only used as guide for modelling):

$$n\lambda = 2d \sin\theta \qquad \text{Equation 1}$$

where n is an integer, λ is the wavelength of incident wave (1.5418 Å for the $Cu_{K\alpha}$ radiation), d is the spacing between the planes in the lattice, and θ is the angle between the incident ray and the scattering planes. The calculated interlayer distances are listed in *Table 11*.

Compound	Selected reflection (2θ°)	d (Å)	d (pm)
E32F	17.94	4.94	494
Z32F	13.63	6.49	649
E33F	18.16	4.88	488
Z33F	19.57	4.53	453
E32T	16.73	5.29	529
Z32T	10.76	8.21	821
E33T	18.91	4.69	469
Z33T	21.99	4.04	404
E32S	16.73	5.29	529
Z32S	20.04	4.43	443

Table 11 Calculated interlayer distances for the five-membered variants

5.5 Self-assembling of the studied molecules over polycrystalline gold and silver surfaces

The *self–assembling* capabilities were studied over 100nm thick polycrystalline *gold* and *silver* surfaces created by the PLD technique (*4.1.3*). Before depositing the organic matter, the metal layers were checked by *scanning electron microscopy* (SEM, *4.2.2*) to gain some information about the morphology of the surface. The results are shown in *Figure 61*.

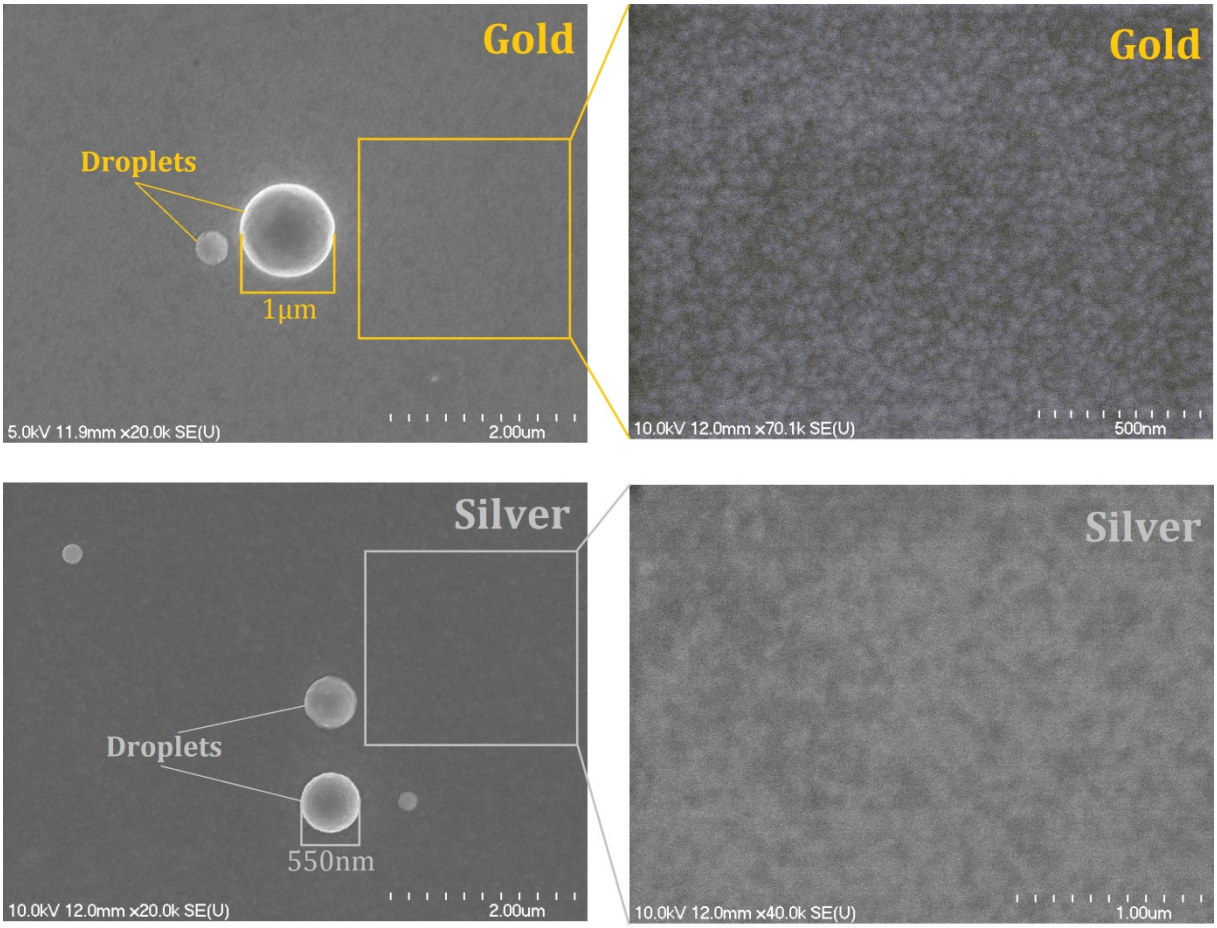


Figure 61 SEM images of the pristine *gold* and *silver* layers made by the PLD technique

As we can see, the layers are enfolded with differently sized droplets of 50–1000 nm (the average size of the droplets is 500 nm, determined by size distribution analysis); however, the spaces between these objects are smooth; therefore, the topology of the *assembled* structures of the deposited molecules are possible to study by *atomic force microscopy* (AFM, *4.2.2*) with relatively high precision (±2 nm deviation).

The metal layers were treaded with the 0.1 M solution in *chloroform* of the studied isomers, in the way mentioned earlier (*4.1.3*), and the presence of the deposited molecules were checked by *infrared microspectroscopy* (IRM, *4.2.2*). Due to the low concentration of the molecules on the metal surface, the main and very intense vibrations could only be identified beside the huge vibrations and rotations of the atmospheric components like *water* and CO_2 (marked on the spectra). Since all other bands were barely higher than the noise level, the spectra were only used for checking the success of the deposition Thus, IRM spectra corresponding to one molecule is shown (*Figure 62*).

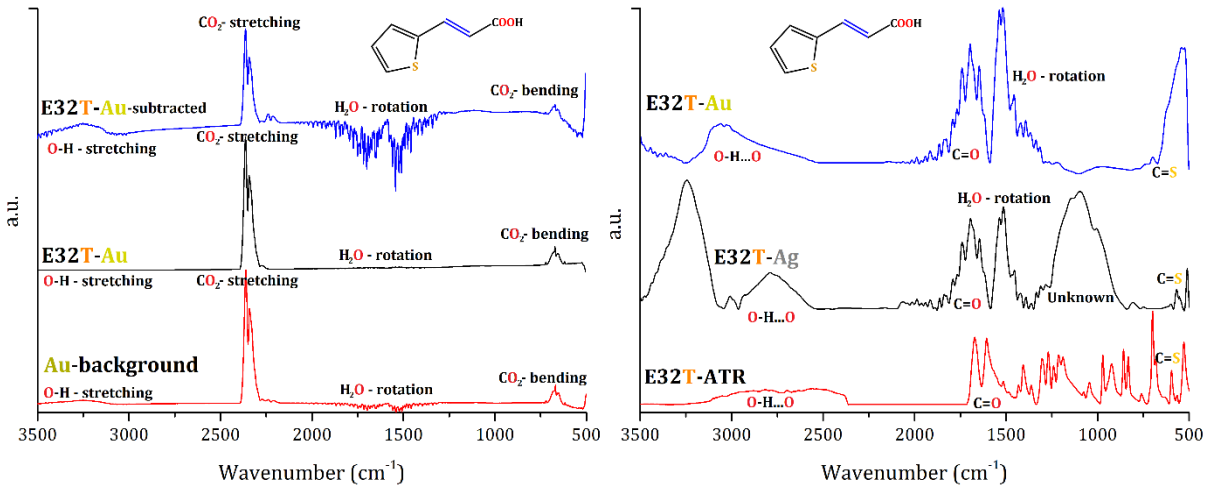


Figure 62 Deposition of the **E32T** molecules on *gold* and *silver* surfaces followed by IRM

A few vibrations definitely corresponding to the organic material can be assigned like the C=S vibrations at 600 cm^{-1} (cannot identified unambiguously in all cases due to the C=O bending), the C=O stretching at 1600 cm^{-1} and the wide O–H…O stretching around 3000 cm^{-1}, which means that the dimers are still kept together via O–H…O=C *hydrogen bond*, i.e., the structure of the absorbed layer may close resemble that of the crystalline state.

After being ascertained about the success of the deposition, the adsorbed layers were studied by AFM. The AFM images of the two different metal surfaces were recorded first. Then, the metal surfaces were treated with the *octylesters* of the two *E-thienyl isomers*, **E32TO** and **E33TO**. (Due to the steric hindrance between the *alkyl* chain and the *thiophene* ring, the orientation of the *Z-thienylcinnamic acid octylesters* cannot be studied this way; nevertheless, because of the structural similarities, the fashion of orientation should be the same as of the *E* isomers. The aim of this investigation was to determine the orientation of the molecules to the surface. The *sulfur* in the *thienyl* group has significantly more affinity to adsorb on the *gold* and *silver* surfaces than either the *ester* group or the *alkyl* chain. Thus, the resulting layer thickness is very small (1–2 nm) if the molecules adsorb parallel to the surface, and becomes relatively large if they do it in a perpendicular fashion. *Figure 63* shows the results of the orientation experiments on the example of **E32TO** molecules.

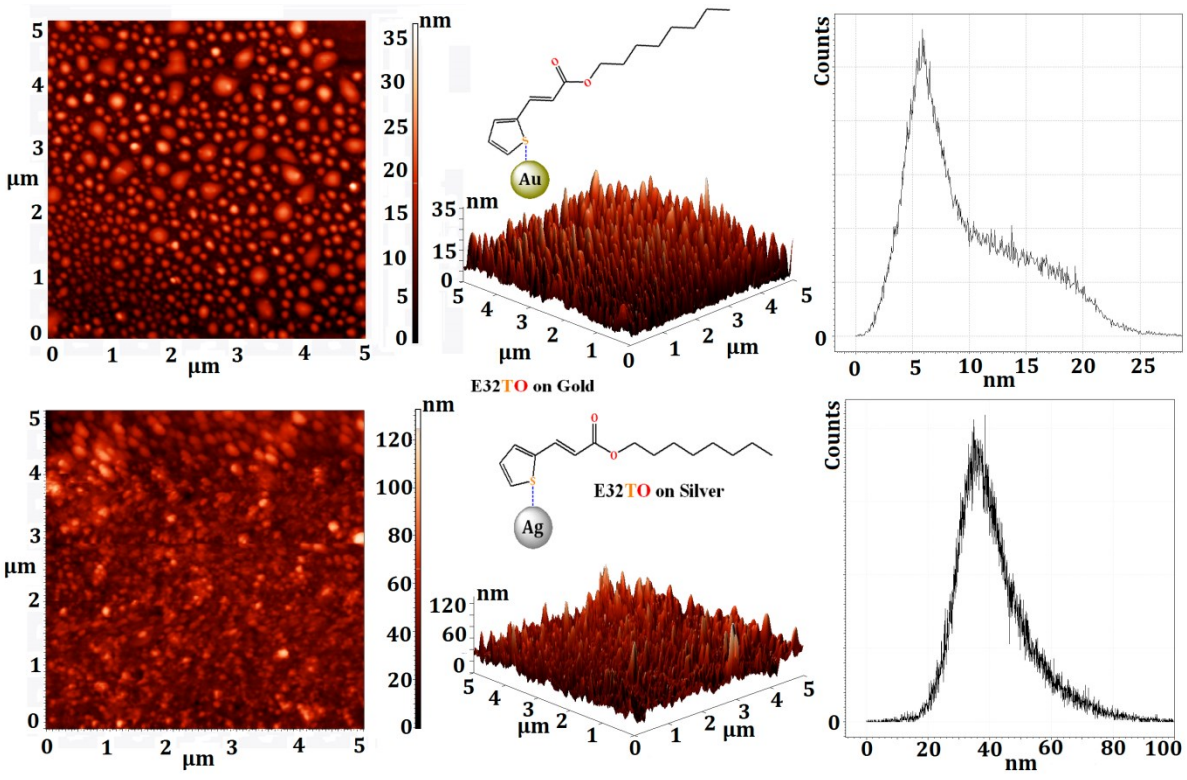


Figure 63 The AFM result of the orientation experiment on both *gold* and *silver* surfaces

It is seen that the coverage is homogenous over both surfaces; however, the thickness of the adsorbed layers differs from each other. Over the *gold* surface, it is close to a single monomolecular carpet; thus, the orientation should be nearly perpendicular. The opposite was found for the *silver* surface, where the adsorbed layer is quite thick, which is a sign of a multilayer formation, which requires nearly parallel orientation of the molecules to the silver surface. Similar results were obtained for the **E33TO** isomer over both surfaces.

Three types of general examples were identified for the different formation types of the studied molecules over polycrystalline *gold* and *silver* surfaces. They are displayed in *Figure 64.*

Type 1: The adsorption of the **E32F** molecule over the *gold* surface leads to homogenous coverage and close to uniform layer thickness. This is the most frequent situation, where the average height of the peaks is 10–38 nm and the average space between them is 30-38 nm (**E32F-Ag**, **E33T-Ag**, **E33F-Ag**). The topology of **E34P** on *silver* is common as well, with heavy organization and lower homogeneity: average high of peaks is 25-45 nm with 100nm inter-peak distance (**E33P-Au**, **E32P-Ag**, **E32P-Au**,

E33F-Au). There are examples (**E33T-Au**, **E32T-Au**), when the number of adsorbed layers is very low (maximum height of peaks is 4 nm), but the coverage is nearly complete (peak to peak distance is <10 nm).

Type 2: Two examples are shown for higher-order organization: **Z22T32T** on *gold* and **Z23P3Ph** on *silver* surface, where the average high of peaks is 25-85 nm and the inter-peak distance is 100 nm.

Type 3: The last examples are not follow the *self-assembling* phenomena, because the formations are derived by surface defects like metal droplets. We can see two different cases; where the molecules are deposited on a "huge" (2 × 2 ×1 μm) metal droplet (**Z23T3Ph-Au**) and the curvature is facilitate the organization. In the other case, where the **Z22T32T** molecules are organized around a "small" (100 × 100 × 50 nm) droplet in a concentric circular fashion (d = 9 μm), governed by the surface defect.

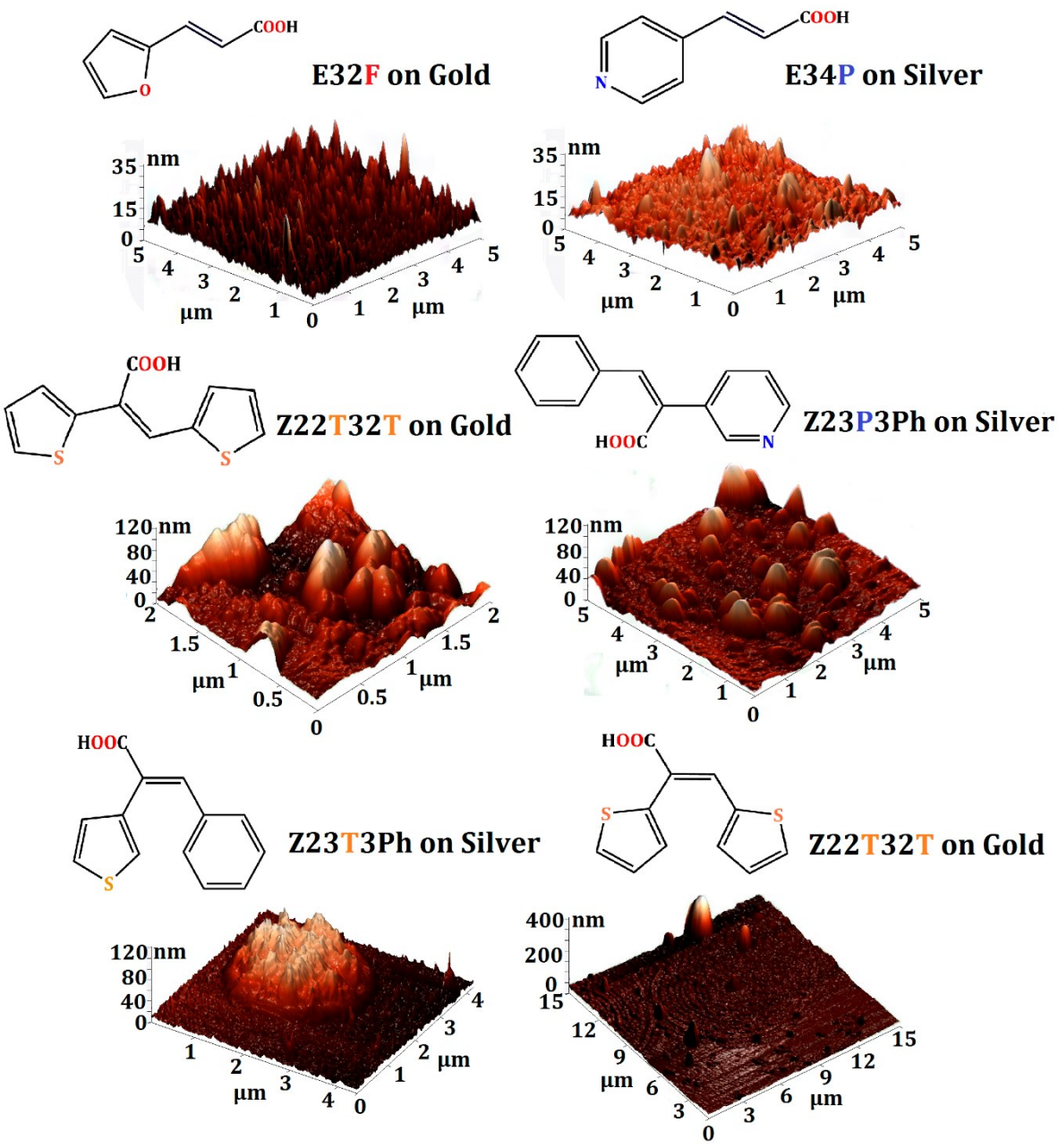


Figure 64 Selected 3D AFM images of the studied materials on both *gold* and *silver* surfaces: **type 1***:* **E32F**–Au, **E34P**–Ag; **type 2**: **Z22T32T**–Au, **Z23P3Ph**–Ag; **type 3**: **Z23T3Ph**–Ag, **Z22T32T**–Au

Based on literature data and our own measurements, we propose that the mechanism of the *self-assembling* is the following: The first step is the covalent adsorption on the metal surface of the dimers (for the *five-membered* derivatives) or the monomers (in the case for the *pyridyl* compounds) via the heteroatom of the aromatic ring. Then, when the surface is saturated with the adsorbed species, the molecular assemblies start to grow, and the components are kept together with weaker interactions, like weak (aromatic)C–H…S(N) *hydrogen bonds* and *π-stacking* interactions.

Due to surface defects, the studied derivatives form multimolecular carpets with different thicknesses, and the assembled molecules stand nearly perpendicular to the surface. The size of the aggregates found on the metal surfaces were varied; however, the estimated number of the molecules in the layer was always more than a single dimer; consequently, it is safe to state that the weak (aromatic)C–H…S(N) *hydrogen bonds* and *π-stacking* interactions play dominant role in the formation of *self-assembled* structures over these metal surfaces. The schematic representation of the resulting structure is given in *Figure 65*.

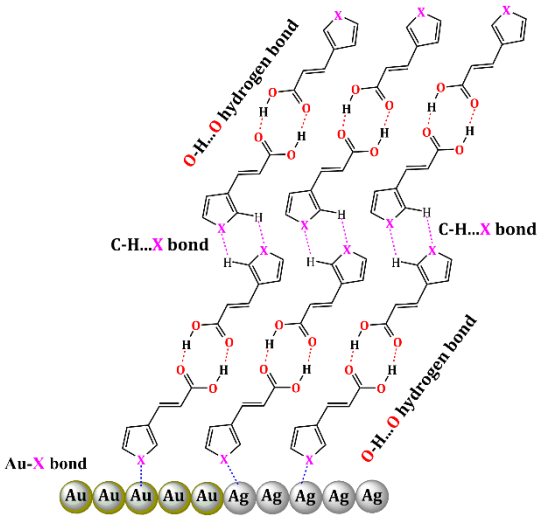


Figure 65 The schematic reperesentation of the *self-assembled* multilayer over Ag and Au surfaces

5.6 Computational results

5.6.1 Conformational analysis

For the conformational study the "random walk" algorithm was chosen, because it covers the potential energy surface well with significantly less calculated points. An example for the results of a conformational analysis is represented with **E32P** (performed at **HF/6-31G(d,p)** level with **Gaussian 09**' and processed with **Izzy Reader v.1.0**) is shown in *Figure 66*. The structure of the potential surface is quite typical for *cinnamic acid* derivatives. The structure with the lowest energy, designated as "A", is located in a potential energy pit in the middle of the potential energy surface. Here, both *pyridyl* and *carboxyl* dihedrals are 0°, which means completely flat molecular structure. The other minima on the surface are marked as "B", "C" and "D". They represent the four other possible conformers. This was the situation for all the molecules studied. The heights of the rotation barriers dividing the minima are in the 3–9 kcal/mol range for

the **E32P** isomer. Another interesting observation is the anomalously high potential barrier at −90°. This is a sign of intramolecular *hydrogen bonds*; however, the molecule can rotate through lower barriers to reach other conformations. Due to the relatively low rotation barriers, one can predict close to free rotation for all rotatable part of the molecules, completely coinciding with the results of the NOESY experiments.

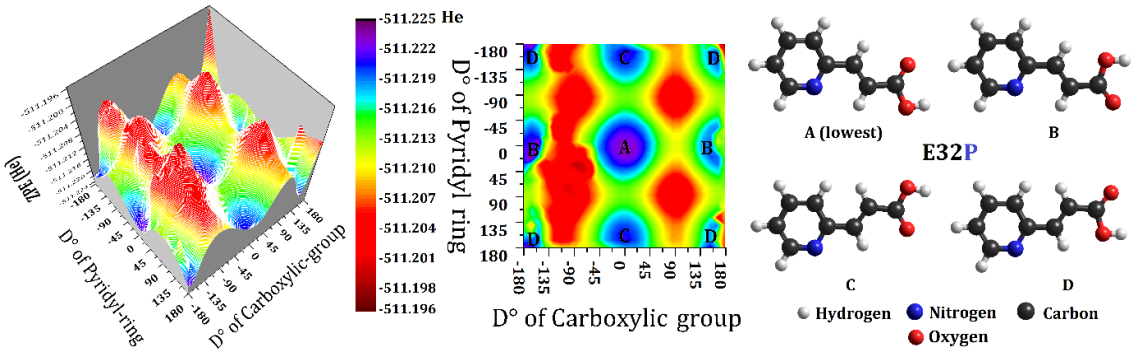


Figure 66 The calculated potetntial energy surface of **E32P** isomer and the resulted four conformers

The potential energy surfaces of the *α-substituted heteroaryl cinnamic acids* are not shown here, because they have three dihedral angles, thus, the potential energy surfaces cannot be represented in 3D diagrams.

The most stable conformers resulting from the conformational analysis of all studied molecules are depicted in *Figure 67*. As has been mentioned already, they are all flat except the *zwitterionic* forms, where the *carboxylate* group turns out of the plane with 48–117°. The **Z32P-ZW** derivative has completely twisted structure due to the close contact of the differently charged parts of the molecule.

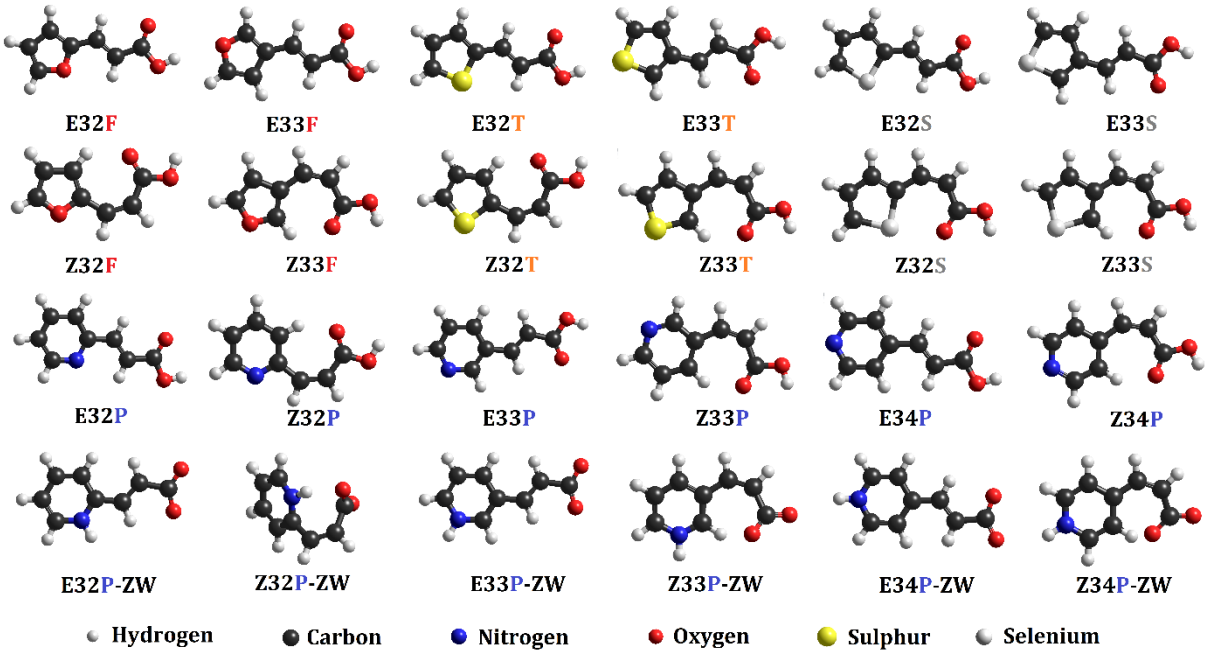


Figure 67 The most stable conformers of the studied molecules calculatd at **HF/6-31G(d,p)** level

The resulting structures from the conformational analysis are used as the starting geometries for the higher level calculations (**B3LYP/CBSB7**, **B3LYP/6-31++G(d,p)**, **WB97XD/CBSB7**, **WB97XD/6-31++G(d,p)**), and the thermally corrected (to 298K) *zero-point energies* (ZPE) and the energy difference between the stereoisomers are listed in *Table 12* for several different DFT methods and basis sets. It is seen that the *E isomer* is the thermodynamically more stable state in each case, as it was indicated by the experimental results in the synthetic part *4.3.2*.

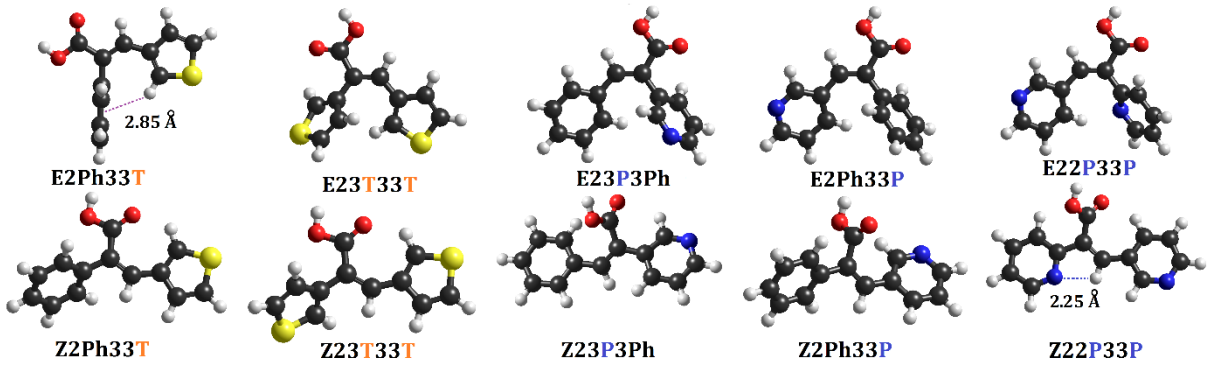


Figure 68 The most stable conformers of the selected *α-substituted* derivatives, calculated at the **PM6** level

Thermally corrected (to 298K) zero-point energies (ZPE_{corr}) from geometry optimization checked with frequency calculations for the monomers, using the lowest structure from conformational analysis as the input.								
Compound	**B3LYP/CBSB7 ZPE_{corr} (He)**	***E–Z* difference (kcal/mol)**	**B3LYP/6-31++G(d,p) ZPE_{corr} (He)**	***E–Z* difference (kcal/mol)**	**WB97XD/CBSB7 ZPE_{corr} (He)**	***E–Z* difference (kcal/mol)**	**WB97XD/6-31++G(d,p) ZPE_{corr} (He)**	***E–Z* difference (kcal/mol)**
E32F	-496.144524	-2.72	-496.042186	-2.76	-495.971373	-2.55	-495.8724561	-2.55
Z32F	-496.140191		-496.037791		-495.967297		-495.8683909	
E33F	-496.138857	-2.44	-496.036746	-2.66	-495.967292	-2.27	-495.8684852	-2.40
Z33F	-496.134967		-496.032507		-495.963674		-495.8646622	
E32T	-819.129919	-3.44	-819.017331	-3.90	-818.961618	-3.47	-818.8521323	-3.87
Z32T	-819.124437		-819.011110		-818.956088		-818.8459634	
E33T	-819.128781	-3.05	-819.016004	-3.37	-818.961627	-2.96	-818.8519007	-3.22
Z33T	-819.123927		-819.010630		-818.956905		-818.8467706	
E32S	-2822.451564	-1.34	-2820.237288	-1.69	-2822.327049	-1.50	-2820.10881	-1.78
Z32S	-2822.449432		-2820.234588		-2822.324650		-2820.105973	
E33S	-2822.449507	-3.40	-2820.234428	-2.87	-2822.326252	-3.30	-2820.10706	-2.71
Z33S	-2822.444087		-2820.229859		-2822.320979		-2820.102741	
E32P	-514.402659	-4.69	-514.301571	-4.78	-514.219814	-4.73	-514.1223717	-4.77
Z32P	-514.395185		-514.293951		-514.212263		-514.1147685	
E33P	-514.399025	-3.48	-514.298383	-3.66	-514.216424	-3.59	-514.1191916	-3.70
Z33P	-514.393472		-514.292549		-514.210690		-514.1132911	
E34P	-514.398239	-4.46	-514.297655	-4.65	-514.215743	-4.59	-514.11862	-4.76
Z34P	-514.391136		-514.290245		-514.208416		-514.1110311	

Table 12 Energy results for the optimized monomers of *heteroaryl cinnamic acids* using several different basis sets and functional combinations

Thermally corrected (to 298K) zero-point energies (ZPE_{corr}) from geometry optimization checked with frequency calculations for the monomers of α-substituted *heteroaryl cinnamic acids*, using the lowest structure from conformational analysis as the input.		
Compound	**B3LYP/CBSB7 ZPE_{corr} (He)**	***E–Z* difference (kcal/mol)**
E2Ph33T	-1050.230284	-2.70
Z2Ph33T	-1050.225975	
E23T33T	-1370.993100	1.84
Z23T33T	-1370.996034	
E23P3Ph	-745.500469	-3.94
Z23P3Ph	-745.494178	
E22P33P	-761.539032	-0.36
Z22P33P	-761.538447	
E2Ph33P	-745.499894	-3.84
Z2Ph33P	-745.493772	

Table 13 Energy results for the optimized monomers of *α-substituted cinnamic acids* at **B3LYP/CBSB7** level

The geometry and energy data derived from the conformational analysis of the *α-substituted* derivatives are shown in *Figure 68* and *Table 13.* The results differ from the unsubstituted *heteroaryl cinnamic acids*, because none of them have planar structures due to the steric hindrance among the bulky groups on the olefinic skeleton. In some cases intramolecular interactions could be identified, e.g. a C–H…N bond (2.25 Å) in **Z22P33P** or a C–H…π interaction (2.85 Å) in **E2Ph33T**. Another important observation is the energy difference between the stereoisomers, where the energies are close to zero or even have positive value, which means among the *α-substituted* derivatives, occasionally, the *Z isomer* is the thermodynamically more stable form in some cases. This finding correlates well with the experimental observations, e.g. with the isomer distribution in the modified *Perkin* condensation.

5.6.2 Hydrogen-bonded dimers

As we have seen already, the fundamental unit of the *self-assembled* structures is the cyclic double *hydrogen-bonded* dimers kept together with strong O–H…O *hydrogen bonds* between the *carboxylic* groups. These dimers were constructed from the optimized monomers for further calculations. Some examples for the *hydrogen-bonded* dimers are shown in *Figure 69*. The other derivatives looks completely similar: the dimeric forms of each *heteroaryl cinnamic acid* derivative, both *E* and *Z isomers* have completely planar structure as it is indicated in *Table 14*. The table shows the energy and geometry data for the *hydrogen-bonded* dimers calculated using two different basis

sets with the **B3LYP** functional (**B3LYP/CBSB7** and **B3LYP/6-31++G(d,p)**). The bond lengths of the cyclic hydrogen bonds are quite regular for each derivative: 1.67 ± 0.02 Å with the **CBSB7** basis set and 1.64 ± 0.02 Å with the 6-31++G(d,p). The bonding energies look similar as well, and they are independent from the type of the *heterocycle*. However, there are significant differences between the *E* and *Z isomers*; the bonding energy is 1 kcal/mol smaller than that in the *E dimers*. The average bonding energy is 14.6 ± 0.7 kcal/mol with the **CBSB7** set and 15.1 ± 0.7 kcal/mol for the **6-31++G(d,p)** basis set, which correlate well with the experimental data for the dissociation energy of *carboxylic acids*.

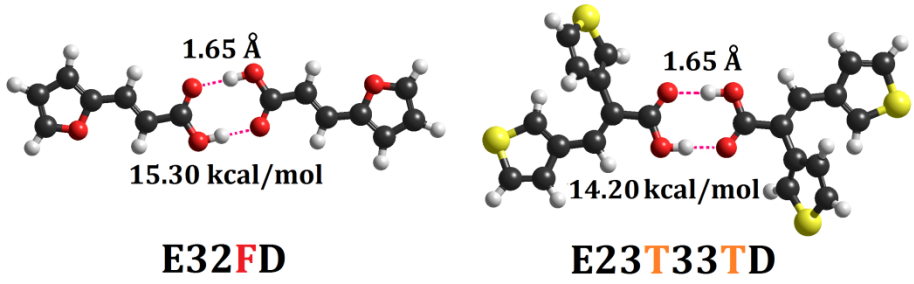


Figure 69 Structures, *hydrogen bond* length and energy data for two *hydrogen-bonded* dimers (**E32FD** and **E23T33TD**) calculated at **B3LYP/CBSB7** level

The **ZPE** of the dimers are **BSSE** corrected, and the bonding energies of the *hydrogen bonds* are calculated using the following formula (it produces positive bonding energy value if it is favorable and negative if it is not):

$$E_{H-bond} = (2 * ZPE_{monomer}) - ZPE_{dimer\,(BSSE_{corrected})} \qquad \text{Equation 2}$$

The *α-substituted* derivatives form more interesting structures, and their geometries are different from those of the corresponding monomers. In *Figure 70* one can see the effect of steric hindrance among the aromatic rings, which makes the shape of the O-H…O=C vibrations different in the infrared spectrum.

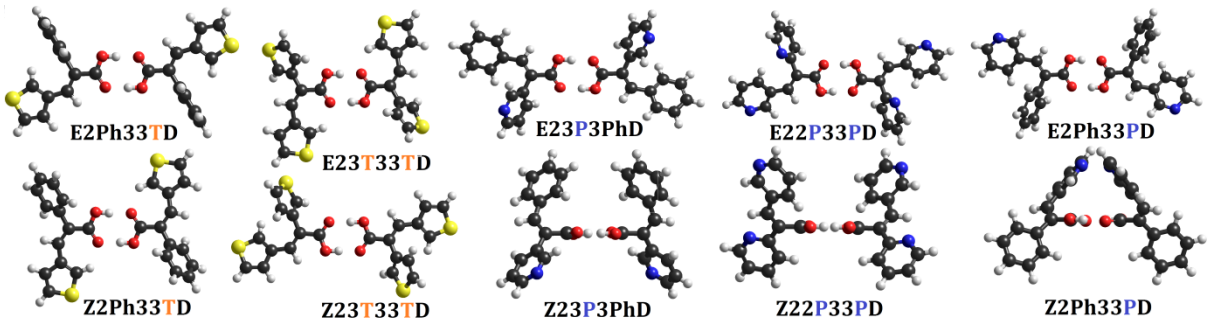


Figure 70 The optimized structures of some *α-substituted heteroaryl cinnamic acid* derivative calculated at **B3LYP/CBSB7** level

Thermally (to 298K) and BSSE corrected zero-point energies (ZPE_{corr}) from geometry optimization checked with frequency calculations for the *hydrogen-bonded* dimers (O−H...O=C) using the optimized monomers as the input.

Compound	B3LYP/CBSB7 ZPE_{corr} (He)	H-bond E_{corr}	$BSSE_{corr}$ (kcal/mol)	H-bond length (Å)	D°	B3LYP/6-31++G(d,p) ZPE_{corr} (He)	H-bond E_{corr}	$BSSE_{corr}$ (kcal/mol)	H-bond length (Å)	D°
E32FD	-992.313422	15.30	3.43	1.65	0.1	-992.1096493	15.86	0.83	1.64	0.10
Z32FD	-992.302794	14.06	3.17	1.68	0.7	-992.0989242	14.65	0.97	1.65	0.80
E33FD	-992.302200	15.37	3.40	1.65	0.0	-992.0988963	15.94	0.82	1.63	0.10
Z33FD	-992.292002	13.85	3.15	1.68	0.0	-992.0879833	14.41	1.00	1.65	0.00
E32TD	-1638.284314	15.36	3.39	1.65	0.9	-1638.060023	15.91	0.84	1.63	0.00
Z32TD	-1638.271129	13.96	3.14	1.68	0.0	-1638.045377	14.53	0.97	1.65	0.00
E33TD	-1638.282083	15.39	3.39	1.65	0.0	-1638.057438	15.96	0.80	1.63	0.10
Z33TD	-1638.269825	13.79	3.13	1.68	0.3	-1638.044095	14.33	0.96	1.66	0.40
E32SD	-5644.927597	15.35	3.38	1.65	0.0	-5640.499972	15.94	1.36	1.62	0.00
Z32SD	-5644.921462	14.18	3.17	1.68	0.0	-5640.492421	14.59	2.62	1.64	0.00
E33SD	-5644.923500	15.37	3.38	1.65	0.4	-5640.494238	15.93	0.87	1.63	0.00
Z33SD	-5644.910203	13.82	3.12	1.68	0.0	-5640.482558	14.33	1.97	1.65	0.00
E32PD	-1028.829406	15.12	3.40	1.66	0.6	-1028.628092	15.66	0.87	1.65	0.00
Z32PD	-1028.812458	13.86	3.09	1.68	0.8	-1028.61082	14.38	0.97	1.66	0.60
E33PD	-1028.822748	15.50	3.30	1.65	0.0	-1028.622334	16.05	0.84	1.65	0.00
Z33PD	-1028.809013	13.85	3.07	1.68	0.0	-1028.608003	14.37	0.96	1.66	0.00
E34PD	-1028.821017	15.40	3.26	1.66	0.8	-1028.620681	15.92	0.84	1.64	0.00
Z34PD	-1028.804456	13.92	3.02	1.68	0.0	-1028.603464	14.42	0.94	1.66	0.30

Table 14 Energy results for the optimized double *hydrogen-bonded* dimers of *heteroary lcinnamic acids* at two different basis sets with **B3LYP** DFT function combination

Thermally (to 298K) and BSSE corrected zero-point energies (ZPE_{corr}) from geometry optimization checked with frequency calculations for the hydrogen-bonded dimers (O–H…O=C) using the optimized monomers as the input.

Compound	B3LYP/CBSB7 ZPE_{corr} (He)	H-bond E_{corr}	$BSSE_{corr}$ (kcal/mol)	H-bond length (Å)	D°
E2Ph33T	-2100.484734	15.16	3.31	1.65	0.01
Z2Ph33T	-2100.473507	13.52	2.95	1.68	0.20
Z23T33T	-2742.007903	13.61	2.94	1.68	3.96
E23T33T	-2742.016754	15.48	3.27	1.65	1.35
E23P3Ph	-1491.025582	15.46	3.23	1.65	0.44
Z23P3Ph	-1491.010583	13.94	3.42	1.67	5.64
E22P33P	-1523.102579	15.38	3.25	1.65	2.30
Z22P33P	-1523.094680	11.15	5.84	1.68	3.50
E2Ph33P	-1491.024651	15.59	3.23	1.65	0.26
Z2Ph33P	-1491.010190	14.20	3.44	1.67	5.30

Table 15 Energy and *hydrogen bond* length data for the optimized double *hydrogen-bonded* dimers of *α-heteroaryl* substituted *cinnamic acids* at **B3LYP/CBSB7** level

Somewhat higher ambiguity was found in the bonding energies of the *α-substituted* derivatives (11.15–15.59 kcal/mol); however, the average bond length is the same as was obtained for the *heteroaryl cinnamic acid* derivatives (1.65–1.68 Å). Consequently, it may be stated that the properties of the double *hydrogen bond* are relatively independent from the structure of the molecules studied molecules, and the formation of this bond is favorable in each case, which correlates well with the results of the infrared experiments, even in the case of the *pyridyl* derivatives.

5.6.3 Hydrogen-bonded tetramers and other structures

The tetramers were constructed from the optimized dimers, and the same basis sets were used to optimize the structures. Unfortunately, the **6-31++G(d,p)** basis set failed to handle the tetramers. However, the **CBSB7** basis set provided with realistic data concerning the geometry and bonding energy data for each isomer. Some derivatives failed to converge with the CBSB7 set as well (indicate with red cross); however, the refined structures did converge resulting in other geometries due to the other possible interactions like the (aromatic)C–H…O=C *hydrogen bond* and various *π-stacking* interactions. In those cases where the convergence leads to other geometries than the *hydrogen-bonded* tetramer, the bond length was fixed to 300 pm for the *thienyl* and 325 pm for the *selenophenyl* derivatives to get precise bonding energies (**BSSE**

correction was applied and the bonding energy was calculated similarly to *Equation 2*). An example for the *hydrogen-bonded* tetramer is shown in *Figure 71*.

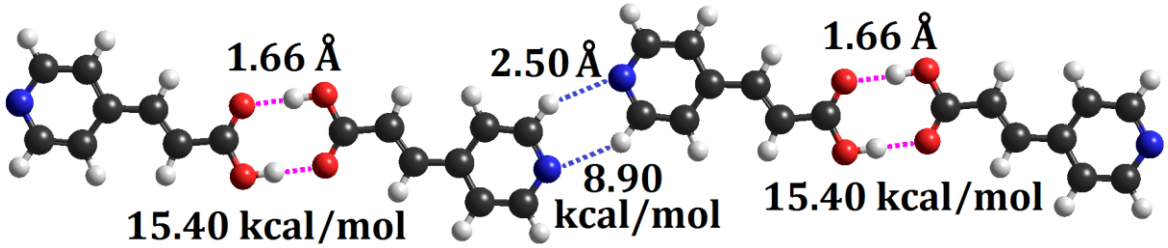


Figure 71 The *hydrogen-bonded* tetramer (**E34PT**) of **E34P** optimized at the **B3LYP/CBSB7** level

The bonding energies and geometry parameters for the other derivatives are listed in *Table 16*. The energies for the (aromatic)C–H…X *hydrogen bonds* varied from 2.63 to 9.10 kcal/mol, and in the order of (aromatic)C–H…N > (aromatic)C–H…O > (aromatic)C–H…S > (aromatic)C–H…Se, which is similar to the results of $^{1}H–^{1}H$ COSY measurements.

Some interesting arrangements, obtained after adjusting the problems with convergence encountered with the *sulphur*- and *selenium*-containing compounds, are depicted in *Figure 72*.

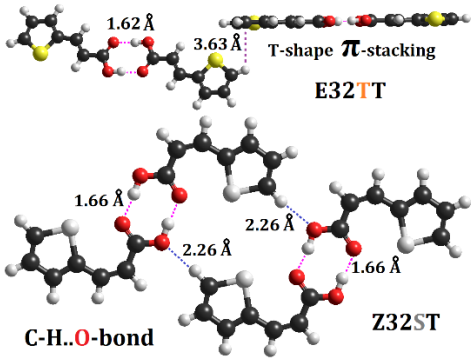


Figure 72 C–H…O-bonded and *π-stacked* geometries from *hydrogen-bonded* tetramer starting structures

The arrangement for **E32TT** is favorable energetically, since the *π-stacking* energy was calculated to be 3.64 kcal/mol (**BSSE** corrected), which is close to that C-H…S bonding energy. The arrangement obtained for **Z32ST** is energetically more favorable (5.02 kcal/mol, **BSSE** corrected) than the simple *H-bonded* version.

Due to the very high computational cost of the *α-substituted* tetramers, geometry optimizations were not performed for the corresponding tetramers.

Thermally (to 298K) and BSSE corrected zero-ooint energies (ZPE_{corr}) from geometry optimization checking with frequency calculations for the hydrogen-bonded tetramers [(aromatic)C–H…X–C(aromatic)] using the optimized dimers as inputs.										
Compound	**B3LYP/CBSB7 ZPE_{corr} (He)**	**H-bond E_{corr} (kcal/mol)**	**$BSSE_{corr}$ (kcal/mol)**	**H-bond length (Å)**	**D°**	**B3LYP/6-31++G(d,p) ZPE_{corr} (He)**	**H-bond E_{corr} (kcal/mol)**	**$BSSE_{corr}$ (kcal/mol)**	**H-bond length (Å)**	**D°**
E32FT	-1984.639383	7.87	7.78	2.53	1.00	-1984.222065	1.74	0.44	2.53	1.00
Z32FT	-1984.616973	7.14	7.34	2.57	0.07	-1984.200984	1.97	0.36	2.57	0.09
E33FT	-1984.616798	7.78	7.92	2.56	2.30	-1984.200594	1.76	0.35	2.56	4.53
Z33FT	-1984.595946	7.49	7.47	2.56	0.03	-1984.179941	2.49	0.35	2.56	0.02
E32TT	-3276.575778	4.49	7.22	3.00	0.00	-3276.119813	-0.15	0.34	3.00	0.00
Z32TT	-3276.548557	3.95	6.55	3.00	0.00	-3276.088814	-1.22	0.31	3.00	0.00
E33TT	-3276.572734	5.38	6.55	3.00	0.00	-3276.114264	-0.38	0.29	3.00	0.00
Z33TT	-3276.547281	4.79	6.46	3.00	0.00	-3276.087813	-0.24	0.30	3.00	0.00
E32ST	-11289.85939	2.63	7.18	3.25	0.00	-11280.99572	-2.65	2.80	3.25	0.00
Z32ST	-11289.85055	4.79	7.28	3.25	0.00	-11280.99213	4.57	4.37	3.25	0.00
E33ST	-11289.85471	4.84	7.20	3.25	0.00	-11280.98673	-1.10	3.41	3.25	0.00
Z33ST	-11289.82577	3.37	6.57	3.25	0.00	-11280.96711	1.25	2.09	3.25	0.00
E32PT	-2057.662339	8.14	8.14	2.53	0.00	-2057.252594	3.13	0.41	2.53	0.00
Z32PT	-2057.636666	7.37	7.57	2.54	0.00	-2057.226827	3.26	0.32	2.54	0.00
E33PT	-2057.660002	9.10	8.14	2.50	0.00	-2057.249931	2.51	0.38	2.51	0.01
Z33PT	-2057.620182	7.95	7.95	2.51	0.00	-2057.210314	3.57	0.35	2.51	0.00
E34PT	-2057.656222	8.90	8.07	2.50	0.03	-2057.246441	2.86	0.34	2.50	0.02
Z34PT	-2057.622386	8.46	7.60	2.49	0.00	-2057.212465	3.47	0.31	2.49	0.00

Table 16 Energy and geometry data for the optimized *hydrogen-bonded* tetramers using two different basis sets with the **B3LYP** functional

5.6.1 Π-stacking interactions

The "sandwich" type *π-stacking* was the final intermolecular interaction studied. The optimized monomers (**WB97XD/CBSB7** and **WB97XD/6-31++G(d,p)**) were used to construct the starting geometries, and the selected distances among the centers of the aromatic rings was 3.9 Å, based on literature data. One example for the structure is presented with **E34P-π** in *Figure 73*.

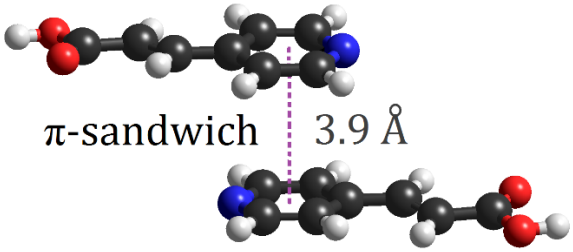


Figure 73 „Sandwich"-type *π-stacking* interaction presented with **E34P** molecules

The resulting bonding energies were calculated similarly to *Equation 2*, (**BSSE** corrected as well), they correlate well with the literature data (calculated for benzene); the difference was about 1.5 kcal/mol. The resulting energies with two different basis sets are listed in *Table 17*.

Similar calculations could not be performed for the α-substituted variants, because of geometric reasons. The molecules are too congested; consequently, proper starting geometries could not be constructed.

Thermally (to 298K) and BSSE corrected zero-point energies (ZPE_{corr}) from geometry optimization checked with frequency calculations for "sandwich" type *π-stacked* dimer using the optimized monomers as inputs.						
Compound	**WB97XD/CBSB7 ZPE_{corr} (He)**	**π-stacking E_{corr} (kcal/mol)**	**$BSSE_{corr}$ (kcal/mol)**	**WB97XD/6-31++G(d,p) ZPE_{corr} (He)**	**π-stacking E_{corr} (kcal/mol)**	**$BSSE_{corr}$ (kcal/mol)**
E32F-π	-991.94894	3.88	0.81	-991.7505443	3.53	0.62
Z32F-π	-991.9406	3.77	0.90	-991.7422873	3.45	0.67
E33F-π	-991.9392	2.90	0.65	-991.740967	2.51	0.59
Z33F-π	-991.93291	3.49	0.76	-991.7345347	3.27	0.73
E32T-π	-1637.9305	4.58	1.23	-1637.711069	4.27	0.82
Z32T-π	-1637.9183	3.86	1.12	-1637.697621	3.57	0.86
E33T-π	-1637.9302	4.33	1.10	-1637.710048	3.92	0.83
Z33T-π	-1637.9211	4.59	1.40	-1637.70047	4.35	1.02
E32S-π	-5644.6608	4.19	0.77	-5640.223655	3.79	8.84
Z32S-π	-5644.6557	4.02	0.83	-5640.217944	3.76	8.21
E33S-π	-5644.6595	4.40	0.80	-5640.220386	3.93	9.65
Z33S-π	-5644.6497	4.87	0.91	-5640.212855	4.63	8.52
E32P-π	-1028.4473	4.80	0.73	-1028.251805	4.43	0.66
Z32P-π	-1028.4315	4.35	0.79	-1028.235965	4.03	0.62
E33P-π	-1028.4399	4.45	0.70	-1028.244973	4.14	0.62
Z33P-π	-1028.427	3.54	0.74	-1028.231579	3.14	0.63
E34P-π	-1028.4364	3.11	0.60	-1028.241734	2.82	0.56
Z34P-π	-1028.4238	4.39	0.83	-1028.228775	4.21	0.74

Table 17 Energy and geometry data for the optimized *π-stacked* dimers using two different basis sets with the **WB97XD** functional

SUMMARY

The family of *cinnamic acids* (*E-3-phenylpropenoic acids*) and they derivatives are discovered 140 years ago; however, they are still intensively studied materials due to their wide-ranging application possibilities (UV-protection, food preservation, producing chemical intermediates and pharmacological products, etc.). In spite of the mentioned applications, one can only find a few literature examples for the preparation, characterization and practical uses of the *heterocyclic* derivatives and the *Z isomers* of *cinnamic acids* (for both *3-arylpropenoic* and *2,3-arylpropenoic acids*); therefore studying these compounds is still relevant.

During the experimental work leading to this dissertation, a variety of *E* and *Z cinnamic acids* were synthesized, including *furyl-*, *thienyl-*, *selenophenyl-*, and *pyridylcinnamic acids* and some *α-substituted* variations containing *thienyl* or *pyridyl* groups, applying numerous types of condensation reactions like the *Perkin-*, *Kövenagel-Döbner-* (leading to the thermodynamically more favorable *E isomer*) and several, modified *Horner-Wadsworth-Emmons* reactions (HWE) (which can be both *E-* or *Z-selective* depending on the reagent and the circumstances). The synthesized materials were extensively characterized by various methods including several spectroscopic techniques, like *infrared spectroscopy* (IR), *nuclear magnetic spectroscopy* (NMR) with numerous 1D and 2D techniques and other analytical tools, e.g. *X-ray diffractometry* (XRD), *atomic force microscopy* (AFM) and *infrared microspectroscopy* (IRM). The experimental findings were checked by quantum chemical calculations performed at various theoretical levels (**PM6**, **HF/6-31G(d,p)**, **B3LYP/CBSB7**, **WB97XD/CBSB7**).

For the synthesis of stereochemically pure *E-3-heteroarylcinnamic acids*, the *Knövenagel-Döbner* condensation with slight modifications (mostly in the reaction time, temperatures and the processing of the reaction mixture) was found to be a general method providing impressive yields and selectivity (100% *E-selectivity* and >95% yields were observed). It was found that the *E-α-substituted* variants can be prepared via both modified *Perkin* and traditional HWE synthesis; however, the HWE technique performed much better (selectivity could be adjusted and the yields were >90%) than the *Perkin* reaction. The hardly accessible *Z-3-heteroarylcinnamic acids* were synthesized as well applying the modified HWE-reaction, where the reagent was the *Ando-phosphonate*. Due to the stabilized intermediate (under cryogenic conditions and inert atmosphere) 90:10

(*Z:E*) selectivity was achieved in most cases, and the yields were also good (>90%). The UV-induced photoisomerization was found to be an affordable and simple method for the preparation of the *Z*-isomers, on the other hand the separation of the isomer mixtures, requires special equipment.

The structure forming properties of the synthesized derivatives were studied with the combination of various instrumental methods in the solid, the liquid (in solution) and in the adsorbed phase as well. Infrared spectroscopy revealed that both *E* and *Z-cinnamic acids* (*3-* and *2,3-arylpropenoic acids*) derivatives containing a *phenyl-* or a *five-membered ring,* form cyclic double *hydrogen bonds* via the *carboxylic* groups in the solid state and in solution as well, if the solvent is aprotic (*chloroform* or DMSO). Only the *pyridyl* variants act differently, where the fundamental structure in the crystalline phase was the *zwitterion*. Nevertheless, it was confirmed that in DMSO, the formation of the *hydrogen-bonded* dimers were also possible. Comparing the infrared results, (registered in different phases) the presence of possible weaker interactions, like C–H…X(O, S, N, Se) *hydrogen bonds* among the aromatic *protons* and the *heteroatom*, or the olefinic *proton* and the *carboxylic oxygen*, are very feasible, because significant shifts in the aromatic vibrations could be identified in the spectra.

The results of the $^{1}H–^{1}H$ *correlation spectroscopy* (COSY) in saturated solution in DMSO, prove the presence of the previous interactions, especially the one between the aromatic rings, where a variety of intermolecular *hydrogen bonds* appeared as off-diagonal cross-peaks, with weak coupling between the non-neighboring *protons*. Another sign of *hydrogen bonds* was found as well: the signal broadening of the aromatic *proton*s (on the ^{1}H spectra), which was significantly larger compared to olefinic *protons*. These findings proved that the olefinic protons were not involved in the extended *hydrogen-bonded* systems in solution. $^{1}H–^{1}H$ *nuclear Overhauser effect spectroscopy* (NOESY) was found to be useful in gaining some information about the conformational behavior of the materials studied. The results show that the aromatic ring of the *3-heteroaryl* derivatives, rotate freely in dilute (0.05 M in DMSO) solution, on the other hand, the result for the *2,3-variants* was quite different. The ring on the second *carbon* rotates relatively freely, but the conformational space of the other one on the third *carbon* was slightly hindered.

The *self-assembling* capabilities were further investigated over polycrystalline metal surfaces (*gold* and *silver*, created with the PLD method) applying the *dip-coating* method in 0.1 M solution in *chloroform*. The created molecular layers were studied with the combination of IRM and AFM techniques. The results of the IRM measurements confirmed that the adsorption of the molecules on the metal surfaces was successful. Further imaging with AFM revealed some interesting, highly organized formations, where the layer thickness was far larger than monomolecular. The presence of the already mentioned *hydrogen bonds* was found to be predominant. Due to the very high peaks on the surface *π-stacking* was also assumed.

The results of the conformational analysis revealed the complete potential energy surfaces of the *heteroaryl cinnamic acid* derivatives and the stable conformers of the *α-substituted* variants. The properties of the conformational space showed consistency with the NOESY experiments. The most stable conformers (determined by the conformational analysis) were further optimized at higher levels, and completely planar structure was found for the *heteroaryl cinnamic acid* and twisted geometries for the other derivatives. Geometry and energy data for the *hydrogen bonded* dimers proved that the fundamental unit was the dimer with 15 kcal/mol average bonding energy. This finding correlates well with the IR measurements. The possible energy order of the weaker *hydrogen bonds* and *π-stacking* interactions were determined successfully as well: C–H…N > C–H…O > C–H…S ~ π–π > C–H…Se, which is very similar to the experimental findings.

The above-listed results clearly attest that *heteroaryl cinnamic acids* are versatile materials. They can form various, extended *hydrogen-bonded* structures in any studied phases. Among many already mentioned applications, the derivatives studied here can have applications in the field of electronics and can act as anticorrosion agents due to their behavior over metal surfaces. Furthermore, they may have pharmacological effects making them interesting target for the medical industry as well.

RELATED PUBLICATIONS

[1] Csankó, K., Darányi, M., Kozma, G., Kukovecz, Á., Kónya, Z., Sipos, P. Pálinkó, I.:
Self-assembling of *Z*-α-pyridylcinnamic acid dimers over polycrystalline Ag and Au surfaces followed by FT-IR and atomic force microscopies,
J. Mol. Struct. 993, 67–72 (2011).
IF: 1.551

[2] Csankó, K., Illés, L., Felföldi, K., Kiss, J.T., Sipos, P., Pálinkó, I.:
C-H…S hydrogen bonds as the organising force in 2,3-thienyl- and phenyl- or 2,3-dithienyl-substituted propenoic acid aggregates studied by the combination of FT–IR spectroscopy and computations,
J. Mol. Struct. 993, 259–263 (2011).
IF: 1.551

[3] Csankó, K., Forgo P, Boros, K., Hohmann, J., Sipos, P., Pálinkó I.,:
Conformational properties of a pyridylsubstituted cinnamic acid studied by NMR measurements and computations,
J. Mol. Struct, 1044, 286–289 (2013).
IF: 1.599

[4] Csankó, K., Kozma, G., Valkai, L., Kukovecz, Á., Kónya, Z., Sipos, P., Pálinkó, I.:
Self-assembling of 2,3-phenyl/thienyl-substituted acrylic acids over polycrystalline gold,
J. Mol. Struct, 1044, 32–38 (2013).
IF: 1.599

[5] Csankó, K., Ruusuvuori, K.I., Tolnai, B., Sipos, P., Berkesi, O., Pálinkó, I.:

Structural features of pyridylcinnamic acid dimers and their extended hydrogen-bonded aggregations
J. Mol. Struct., in press, doi:10.1016/j.molstruc.2014.11.041, (2014)
IF: 1.599

OTHER PUBLICATIONS

[1] Csankó, K., Sipos, P., Pálinkó, I.: Monometallic supported gold catalysts in organic transformations: Ring making and ring breaking,

Catalysts **2**, 101-120 (2012).

IF: –

REFERENCES

[1]. Köhler, F. E.; Wilhelm, B.; Gürke, M., *Medizinal-Pflanzen in naturgetreuen Abbildungen mit kurz erläuterndem* Gera-Untermhaus: Wien, **1887.** Vol. 1, p 77-77.

[2]. Toussaint-Samat, M., *A history of food*. Wiley-Blackwell: Malden, **2009.** p 433-490.

[3]. Paranagama, P. A.; Wimalasena, S.; Jayatilake, G. S.; Jayawardena, A. L.; Senanayake, U. M.; Mubarak, A. M., *J. Natn. Sci. Foundation Sri Lanka* **2001.** *29*, 147-153.

[4]. Letizia, C. S.; Cocchiara, J.; Lapczinsky, A., *Food Chem. Toxicol.* **2005.** *43*, 925-943.

[5]. Hill, L.; Gomes, C.; Taylor, M., *Food. Sci. Tech.* **2013.** *51*, 86-93.

[6]. Pound, A. W.; Pound, J. R., *J. Phys. Chem.* **1933.** *38*, 1045-1049.

[7]. Perkin, W., *J. Chem. Soc.* **1868.** *21*, 181-186.

[8]. Urbach, F., *Biological responses to ultraviolet A radiation*. Valdenmar: Oxford, **1991.**

[9]. Beavis, R. C.; Chait, B. T.; Fales, H. M., *Rapid Commun. Mass Spectrom.* **1989.** *3*, 432-435.

[10]. Kiss, J. T.; Felföldi, K.; Körtvélyesi, T.; Pálinkó, I., *Vibr. Spect.* **2000.** *22*, 63-73.

[11]. Török, B.; Pálinkó, I.; Tasi, G.; Nyerges, L.; Bogár, F., *J. Chromat. A* **1994.** *668*, 353-358.

[12]. Pálinkó, I.; Török, B.; Rózsa-Tarjányi, M.; Kiss, J. T.; Tasi, G., *J. Mol. Struct.* **1995.** *348*, 57-60.

[13]. Kukovecz, Á.; Kiss, J. T.; Pálinkó, I., *J. Mol. Struct.* **1997.** *408–409*, 325-327.

[14]. Kiss, J. T.; Felföldi, K.; Hannus, I.; Pálinkó, I., *J. Mol. Struct.* **2001.** *565–566*, 463-468.

[15]. Ladell, J.; McDonald, T. R. R.; Schmidt, G. M. J., *Acta Cryst.* **1956.** *9*, 195-195.

[16]. Luo, G.; Yang, R.; Lai, X.; Yang, W.; Xie, S.; Zhou, H., *China j. Chinese mater. med.* **1996.** *21*, 744-763.

[17]. Maranz, S.; Wiesman, Z.; Garti, N., *Journal of agricultural and food chemistry* **2003.** *51*, 6268-73.

[18]. Douglas, C.; Hoffmann, H.; Schulz, W.; Hahlbrock, K., *EMBO J.* **1987.** *6*, 1189-1195.

[19]. Peter, K. L.; Corlett, R.; H.T.W., T., *Singapore Biodiversity: An Encyclopedia Of The Natural Environment And Sustainable Development*. Didier Millet: Singapore, **2011.**

[20]. Li, X. W.; Li, J.; Werff, H., Flora of China. Harvard university press: Cambridge, **2013.** pp 38-39.

[21]. Dumas, J. B.; Préligot, E., *Ann. Chim. Physique* **1834.** *57*, 305-334.

[22]. Chiozza, L., *Compt. Rend.* **1854.** *1*, 222-222.

[23]. Promkatkaew, M.; Suramitr, S.; Karpkird, T. M.; Namuangruk, S.; Ehara, M.; Hannongbua, S., *J. Chem. Phys.* **2009.** *131*, 224-306.

[24]. Blake, J.; Walker, G. N.; Fish, M. S., *Org. Synth. Coll.* **1951.** *31*, 35-35.

[25]. Ford, R. A.; Domeyer, B.; Easterday, O., *Regul. Toxicol. Pharmacol.* **2000.** *31*, 166-181.

[26]. Karas, M.; Bachmann, D.; Hillenkamp, F., *Anal. Chem.* **1985.** *57*, 2935-2939.

[27]. Hillenkamp, F.; Karas, M.; Bachmann, D., *Int. J. Mass.* **1987.** *78*, 53-68.

[28]. Hillenkamp, F.; Karas, M., *Anal. Chem.* **1988.** *60*, 2299-2301.

[29]. Beavis, R. C.; Chait, B. T., *Rapid Commun. Mass Spectrom.* **1989.** *3*, 436-439.

[30]. Beavis, R. C.; Chait, B. T.; Chaudhary, T., *Org. Mass Spectrom.* **1992.** *27*, 156-158.

[31]. Hermann, K. M., *Plant Cell* **1995.** *7*, 907-919.

[32]. Rosen, T., In *Comprehensive Organic Synthesis*, Elsevier: **1991.** pp 395-408.

[33]. Dippy, J. F. J.; Evans, R. M., *J. Org. Chem.* **1950.** *15*, 451-456.

[34]. Johnson, J. R., The Perkin Reaction and Related Reactions. In *Organic Reactions*, John Wiley & Sons, Inc.: **2004.**

[35]. Raj, K.; Johnson, J. R., Organic Reaction Mechanism. In *Organic Reactions*, John Wiley & Sons, Inc.: **1998.** pp 199-201.

[36]. Buckles, R. E.; Hausman, E. A., *J. Am. Chem. Soc.* **1948.** *70*, 415-415.

[37]. Claisen, L., *Ber.* **1890.** *23*, 977-977.

[38]. Schmidt, J., *Ber.* **1881.** *14*, 1459-1461.

[39]. Knoevenagel, E., *Ber.* **1898.** *31*, 2596-2619.

[40]. Doebner, O., *Chem. Ber.* **1902.** *35*, 1136-1136.

[41]. Bruckner, R., Organic mechanisms. Springer: **2002.** pp 423-423.

[42]. Wiley, R. H.; Smith, N. R., *J. Am. Chem. Soc.* **1950.** *72*, 5198-5199.

[43]. Kuhn, R.; Badstuber, W.; Grundmann, C., *Chem. Ber.* **1936.** *69*, 98-98.

[44]. Berslow, D. S.; Hauser, C. R., *J. Am. Chem. Soc.* **1940.** *62*, 2385-2388.

[45]. List, B., *Angew. Chem. Int. Ed. Eng.* **2010.** *49*, 1730-1734.

[46]. Heck, R. F.; Nolley, J. P., *J. Org. Chem.* **1972.** *37*, 2320-2322.

[47]. Mizoroki, T.; Mori, K.; Ozaki, A., *Bull. Chem. Soc. Jap* **1971.** *44*, 581-581.

[48]. Tierney, J. P.; Lidström, P., Microwave Assisted Organic Synthesis. Blackwell: Oxford, **2006.** pp 1-21.

[49]. Mobinikhaledi, A.; Foroughifar, N.; Fathinejad, H., *Syn. React. Inorg. Met.* **2008.** *38*, 428-430.

[50]. Gupta, M.; Wakhloo, B. P., *Arkivoc.* **2007.** *1*, 94-98.

[51]. Rohatgi-Mukherjee, K. K., Fundamentals of photochemistry. **1997.** pp 1-218.

[52]. Rao, Y. S.; Filler, R., *Chem. Comm.* **1976.** 471-472.

[53]. Hocking, M. B., *Can. J. Chem.* **1969.** *47*, 4567-4576.

[54]. Khan, M.; Brunklaus, G.; Enkelmann, V., *J. Am. Chem. Soc.* **2008.** *130*, 1741-1748.

[55]. Lindlar, H., *Helv. Chim. Acta* **1952.** *35*, 446-450.

[56]. García-Mota, M.; Gómez-Díaz, J.; Novell-Leruth, G.; Vargas-Fuentes, C.; Bellarosa, L.; Bridier, B.; Pérez-Ramírez, J.; López, N., *Theor Chem Acc* **2011.** *128*, 663-673.

[57]. Urushibara, Y.; Hirota, M., *Nippon Kagaku Zasshi* **1964.** *85*, 679-681.

[58]. Brown, C. A.; Ahuja, V. K., *Chem. Comm.* **1973.** 553-554.

[59]. Wadsworth, W. S., *Org. React.* **1977.** *25*, 73-73.

[60]. Horner, L.; Hoffmann, H.; Wippel, H. G., *Chem. Ber.* **1958.** *91*, 61-63.

[61]. Wadsworth, W. S.; Emmons, W. D., *J. Am. Chem. Soc.* **1961.** *83*, 1733-1738.

[62]. Wittig, G.; Schöllkopf, U., *Chem. Ber.* **1954.** *87*, 1318-1318.

[63]. Wittig, G.; Haag, W., *Chem. Ber.* **1955.** *88*, 1654-1666.

[64]. Michaelis, A.; Kaehne, R., *Ber.* **1898.** *31*, 1048-1055.

[65]. Arbuzov, A. E., *J. Russ. Phys. Chem. Soc.* **1906.** *38*, 687-687.

[66]. Bhattacharya, A. K.; Thyagarajan, G., *Chem. Rev.* **1981.** *81*, 415-430.

[67]. Larsen, R. O.; Aksnes, G., *Phosph. Sulf.* **1983.** *15*, 218-219.

[68]. Still, W. C.; Gennari, C., *Tetrahedron Lett.* **1983.** *24*, 4405-4408.

[69]. Breuer, E.; Bannet, D. M., *Tetrahedron Lett.* **1977.** *18*, 1141-1144.

[70]. Patois, C.; Savignac, P., *Tetrahedron Lett.* **1991.** *32*, 1317-1320.

[71]. Ando, K., *J. Org. Chem.* **1999.** *64*, 8406-8408.

[72]. Hoveyda, A. H.; Zhugralin, A. R., *Nature* **2007.** *450*, 243-251.

[73]. Siau, W. Y.; Zhang, Y.; Zhao, Y., *Top. Curr. Chem.* **2012.** *327*, 33-58.

[74]. Crowe, W. E.; Goldberg, D. R., *J. Am. Chem. Soc.* **1995.** *117*, 5162-5163.

[75]. Buckles, R. E.; Bremer, K., *Org. Synth. Coll.* **1953.** *33*, 70-70.

[76]. Karaminski-Zamola, G.; Fiser-Jakic, L.; Bajic, M., *Heterocycles* **1985.** *23*, 313-316.

[77]. Kharasch, M. S.; Jensen, E. V.; Urry, W. H., *Science* **1945.** *102*, 128-128.

[78]. Cadogan, J. I. G.; Duell, E. G., *J. Chem. Soc.* **1962.** *(Resumed)*, 4154-4163.

[79]. Li, S.; Yuan, W.; Ma, S., *Ang. Chem. Int. Ed. Eng.* **2011.** *50*, 2578-2582.

[80]. Arunan, E.; Desiraju Gautam, R.; Klein, R. A.; Sadlej, J.; Scheiner, S.; Alkorta, I.; Clary David, C.; Crabtree, R. H.; Dannenberg, J. J.; Hobza, P.; Kjaergaard, H. G.; Legon, A. C.; Mennucci, B.; Nesbitt, D. J., *Pure Appl. Chem.* **2011.** *83*, 1637-1637.

[81]. Emsley, J., *Chem. Soc. Rev.* **1980.** *9*, 91-124.

[82]. Tolnai, B.; Kiss, J. T.; Felföldi, K.; Pálinkó, I., *J. Mol. Struct.* **2009.** *924–926*, 27-31.

[83]. Tamasi, G.; Defazio, S.; Chiasserini, L.; Sega, A.; Cini, R., *Inorg. Chim. Acta* **2009.** *362*, 1011-1021.

[84]. Filippov, O. A.; Belkova, N. V.; Epstein, L. M.; Shubina, E. S., *J. Organomet. Chem.* **2013.** *747*, 30-42.

[85]. Custelcean, R.; Jackson, J. E., *Chem. Rev.* **2001.** *101*, 1963-1980.

[86]. Hopkins, J. C.; Dryden, D. M.; Ching, W. Y.; French, R. H.; Parsegian, V. A.; Podgornik, R., *J. Coll. Interf. Sci.* **2014.** *417*, 278-284.

[87]. Hunter, C. A.; Sanders, J. K. M., *J. Am. Chem. Soc.* **1990.** *112*, 5525-5534.

[88]. Sinnokrot, M. O.; Valeev, E. F.; Sherrill, C. D., *J. Am. Chem. Soc.* **2002.** *124*, 10887-10893.

[89]. Metrangolo, P.; Resnati, G., *Chem. Eur. J.* **2001.** *7*, 2511-2519.

[90]. Isaacs, E. D.; Shukla, A.; Platzman, P. M.; Hamann, D. R.; Barbiellini, B.; Tulk, C. A., *Phys. Rev. Lett.* **1999.** *82*, 600-603.

[91]. Yakovchuk, P.; Protozanova, E.; Frank-Kamenetskii, M. D., *Nucleic Acid Res.* **2006.** *34*, 564-574.

[92]. Chermahini, A. N.; Mahdivan, M.; Teimouri, A., *Bull. Korean Chem. Soc.* **2010.** *31*, 941-948.

[93]. Sauren, H.; Winkler, A.; Hess, P., *Chem. Phys. Lett.* **1995.** *239*, 313-319.

[94]. Navarrete, A.; Somanathan, R.; Aguirre, G., *Acta Cryst.* **2013.** *69*, 1519.

[95]. Pálinkó, I., *Acta Cryst.* **1999.** *55*, 216-220.

[96]. Gou, Q.; Feng, G.; Evangelisti, L.; Caminati, W., *Chem. Phys. Lett.* **2014.** *591*, 301-305.

[97]. Kukovecz, Á.; Pálinkó, I., *J. Mol. Struct.* **1999.** *482–483*, 463-467.

[98]. Gorbitz, C. H.; Etter, M. C., *J. Am. Chem. Soc.* **1992.** *114*, 627-631.

[99]. Nagy, P. I., *Comp. Theor. Chem.* **2013.** *1022*, 59-69.

[100]. Hyla-Kryspin, I.; Haufe, G.; Grimme, S., *Chem. Phys.* **2008.** *346*, 224-236.

[101]. Balogh, Z.; Pálinkó, I., *J. Mol. Struct. THEOCHEM* **2003.** *623*, 11-15.

[102]. Nobeli, I.; Price, S. L.; Lommerse, J. P. M.; Taylor, R., *J. Comp. Chem.* **1997.** *18*, 2060-2074.

[103]. Peña Ruiz, T.; Fernández-Gómez, M.; López González, J. J.; Koziol, A. E.; Granadino Roldán, J. M., *J. Mol. Struct.* **2004.** *707*, 33-46.

[104]. Chakrabarti, P.; Chakrabarti, S., *J. Mol. Biol.* **1998.** *284*, 867-873.

[105]. McGaughey, G. B.; Gagné, M.; Rappé, A. K., *J. Biol. Chem.* **1998.** *273*, 15458-15463.

[106]. Cockroft, S. L.; Hunter, C. A.; Lawson, K. R.; Perkins, J.; Urch, C. J., *J. Am. Chem. Soc.* **2005.** *127*, 8594-8595.

[107]. Ringer, A. L.; Sinnokrot, M. O.; Lively, R. P.; Sherrill, C. D., *Chem. Eur. J.* **2006.** *12*, 3821-3828.

[108]. Sinnokrot, M. O.; Sherrill, C. D., *J. Phys. Chem. A* **2004.** *108*, 10200-10207.

[109]. Fernandes, M. A.; Levendis, D. C.; Koning, C. B., *Cryst. Eng.* **2001.** *4*, 215-231.

[110]. Love, J. C.; Estroff, L. A.; Kriebel, J. K.; Nuzzo, R. G.; Whitesides, G. M., *Chem. Rev.* **2005.** *105*, 1103-1170.

[111]. Grönbeck, H.; Curioni, A.; Andreoni, W., *J. Am. Chem. Soc.* **2000.** *122*, 3839-3842.

[112]. Vericat, C.; Vela, M. E.; Benitez, G.; Carro, P.; Salvarezza, R. C., *Chem. Soc. Rev.* **2010.** *39*, 1805-1834.

[113]. Gómez, M. M.; Vara, J. M.; Hernández , J. C.; Salvarezza, R. C.; Arvia, A. J., *J. Electro. Chem.* **1999.** *474*, 74-81.

[114]. Muniz-Miranda, M., *Chem. Phys. Lett.* **2006.** *417*, 330-333.

[115]. Matsuura, T.; Shimoyama, Y., *Eur. Phys. J. E* **2002.** *7*, 233-240.

[116]. Poppenberg, J.; Richter, S.; Darlatt, E.; Traulsen, C. H. H.; Min, H.; Unger, W. E. S.; Schalley, C. A., *Surf. Sci.* **2012.** *606*, 367-377.

[117]. Baše, T.; Bastl, Z.; Havránek, V.; Lang, K.; Bould, J.; Londesborough, M. G. S.; Macháček, J.; Plešek, J., *Surf. Coat. Tech.* **2010.** *204*, 2639-2646.

[118]. Huang, I. Y.; Lee, M. C., *Sensor. Actuat. B-Chem.* **2008.** *132*, 340-348.

[119]. Witte, G.; Wöll, C., *J. Mater. Res.* **2004.** *19*, 1889-1916.

[120]. Gooding, J. J.; Mearns, F.; Yang, W.; Liu, J., *Electroanalysis* **2003.** *15*, 81-96.

[121]. Iler, R. K., *J. Coll. Interf. Sci.* **1966.** *21*, 569-594.

[122]. Decher, G.; Hong, J. D., *Makromol. Chem. M. Symp.* **1991.** *46*, 321-327.

[123]. Kotov, N. A., *Nanostruct. Mater.* **1999.** *12*, 789-796.

[124]. Lin, Z. P.; Guo, B. B., *Guangpu Shiyanshi* **2012.** *29*, 1786-1788.

[125]. Pastoriza-Santos, I.; Liz-Marzán, L. M., *Pure appl. chem.* **2000.** *72*, 83-90.

[126]. AIST 1H NMR (E-cinnamic acid).**2014.**

[127]. Marvel, C. S.; King, W. B., *Org. Synth. Coll.* **1929.** *1*, 252.

[128]. Chandrasekhar, S.; Reddy, G. P. K.; Nagesh, C.; Reddy, C. R., *Tetrahedron Lett.* **2007.** *48*, 1269-1271.

[129]. Dean, E. W.; Stark, D. D., *J. Ind. Eng. Chem.* **1920.** *12*, 486-490.

[130]. Allen, C. F. H.; Spangler, W. F., *Org. Synth. Coll.* **1955.** *3*, 377-377.

[131]. Zhanglin, Z.; Lilan, S.; Yaozeng, H., *Synt. Comm.* **1991.** *21*, 1027-1037.

[132]. Rahman, M. A. F. M.; Jahng, Y., *Eur. J. Org. Chem.* **2007.** 379-383.

[133]. Sigma-Aldrich 1H NMR (E-3-(2-thienl)acrylic acid).**2014.**

[134]. Sigma-Aldrich 1H NMR (E-3-(3-pyridyl)acrylic acid).**2014.**

[135]. Bio-Rad 1H NMR (NQ_05356).**2005.**

[136]. Abott, T. W.; Althousen, D., *Org. Synth. Coll.* **1932.** *2*, 270-270.

[137]. Abott, T. W., *Org. Synth. Coll.* **1932.** *2*, 515-515.

[138]. Sigma-Aldrich 1H NMR (3-phenylpropiolic acid).**2014.**

[139]. Kurita, T.; Aoki, F.; Mizumoto, T.; Maejima, T.; Esaki, H.; Maegawa, T.; Monguchi, Y.; Sajiki, H., *Chem. Eur. J.* **2008.** *14*, 3371-3379.

[140]. Lindlar, H.; Dubuis, R., *Org. Synth. Coll.* **1973.** *5*, 880-880.

[141]. Trost, B. M.; Hirano, K., *Org. Lett.* **2012.** *14*, 2446-2449.

[142]. Scapens, D.; Adams, H.; Johnson, T. R.; Mann, B. E.; Sawle, P.; Aqil, R.; Perrior, T.; Motterlini, R., *Dalton Trans.* **2007.** *43*, 4962-4973.

[143]. Yao, Q.; Levchik, S., *Tetrahedron Lett.* **2006.** *47*, 277-281.

[144]. Messik, F.; Oberthür, M., *Synthesis* **2013.** *45*, 167-170.

[145]. Richardson, R. M. New synthesis and reactions of phosphonates. University of Iowa, **2012.**

[146]. Carpino L.A.; McAdams, L. V., *Org. Synth. Coll.* **1970.** *50*, 31-31.

[147]. Thompson, M. *Arguslab*, 4.01; Planaria Software LLC: **2005**.

[148]. Hyperchem, T. *Hyperchem Professional 8.0.6*, 8.0.6; Hypercube Inc. 1115 NW 4th Street, Gainesville, Florida 32601, USA: **2007**.

[149]. J. J. P. Stewart, *J. Comp. Chem.* **1989.** *10*, 209-20.

[150]. Roothaan, C. C. J., *Rev. Mod. Phys.* **1951.** *23*, 69-69.

[151]. Collins, J. B.; Schleyer, P. R.; Binkley, J. S.; Pople, J. A., *J. Chem. Phys.* **1976.** *64*, 5142-51.

[152]. Gaussian, I. *Gaussain 09, Revision D.01*, Gaussian, Inc., Wallingford CT: **2013**.

[153]. IRC, g. *Avogadro*, 1.1.1; **2013**.

[154]. Stewart, J. J., *J. Mol. Model.* **2007.** *13*, 1173-213.

[155]. Ditchfield, R.; Hehre, W. J.; Pople, J. A., *J. Chem. Phys.* **1971.** *54*, 724-724.

[156]. Becke, A. D., *J. Chem. Phys.* **1993.** *98*, 5648-5652.

[157]. Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. R., *J. Comp. Chem. 1983*, 294-301.

[158]. J. A. Montgomery, J. A.; Frisch, J. M. J.; Ochterski, J. W.; Petersson, G. A., *J. Chem. Phys.* **1999.** *110*, 2822-2827.

[159]. Head-Gordon M.; Pople, J. A.; Frisch, M. J., *Chem. Phys. Lett.* **1988.** *153*, 503-506.

[160]. Chai, J. D.; Head-Gordon, M., *PCCP* **2008.** *10*, 6615-6620.

[161]. Simon, S.; Duran, M.; Dannenberg, J. J., *J. Chem. Phys.* **1996.** *105*, 11024-11032.

[162]. Boys, S. F.; Bernardi, F., *Mol. Phys.* **1970.** *19*, 553.

[163]. Bondi, A., *J. Phys. Chem.* **1964.** *68*, 441.

MAGYAR NYELVŰ ÖSSZEFOGLALÓ (HUNGARIAN SUMMARY)

A *fahéjsavat* (*E-3-fenilpropénsav*) és származékait több mint 140 évvel ezelőtt fedezték fel, ennek ellenére a mai napig intenzíven kutatott vegyületek, köszönhetően számos felhasználási lehetőségüknek (mint az UV védelem, élelmiszerek tartósítása és ízesítése, valamint vegyipari és gyógyszeripari köztitermékek előállítása, stb.). Az említett alkalmazási lehetőségek ellenére, a heterociklusos származékaik és a megfelelő *Z* izomerek (beleértve a *3-arilpropénsavakat* és a *2,3-arilpropénsavakat*) előállítására és azok alkalmazási lehetőségeire, csak alig néhány irodalmi példát találunk, így ezen vegyületek tanulmányozása még mindig időszerűnek mondható.

A jelen disszertációhoz vezető kísérleti munka során, számos *E* és *Z fahéjsavat* szintetizáltam, többek közt *furil-*, *tienil-*, *szelenofenil-* és *piridil fahéjsavakat*, valamint néhány *α-szubsztituált változatot*, (melyek *tienil-, vagy piridil* gyűrűt tartalmaztak) több kondenzációs reakciót felhasználva, mint a *Perkin-*, *Kövenagel-Döbner-* (melyekkel a termodinamikailag stabilabb *E* izomer nyerhető) és különböző, *Horner-Wadsworth-Emmons* reakciókat (HWE) (ami *E-* vagy *Z*-szelektív is lehet a reaktánstól és körülményektől függően). Az előállított vegyületeket precíz szerkezetvizsgálatnak vetettem alá, számos spektroszkópiai módszert alkalmazva, mint az *infravörös spektroszkópia* (IR), a *mágneses magrezonancia spektroszkópia* (NMR) több 1D és 2D technika felhasználásával és más analitikai eszközöket is, például *Röntgen diffractometriát* (XRD), *atomi erő mikroszkópiát* (AFM) és *infravörös mikrospektroszkópiát* (IRM). A kísérleti eredményeket kvantum kémiai módszerekkel is modelleztem, több különböző elméleti szinten (**PM6**, **HF/6-31G(d,p)**, **B3LYP/CBSB7**, **WB97XD/CBSB7**).

A sztereokémiailag tiszta *E-3-heteroaromás propénsavak* előállítására, a *Knövenagel-Döbner* kondenzáció módosított változata (ahol elsősorban a reakció hőmérséklete, ideje és a termékek feldolgozása változik) bizonyult a legalkalmasabbnak, kiemelkedően magas termeléssel és szelektivitással (100% *E*-szelektivitás és >95% gyakorlati kitermelés). Az *E-α-szubsztituált fahéjsavak* előállítására, mind a módosított *Perkin*, mind a HWE szintézis alkalmasnak bizonyult, azonban a HWE eljárást sokkal eredményesebbnek találtam (a reakció szelektivitása változtatható és a termelés is viszonylag magas >90%) mint a *Perkin* reakciót. A nehezen hozzáférhető *Z-3-heteroaromás fahéjsavak* szintén előállíthatóak a HWE reakcióval, az *Ando-foszfonátot*

alkalmazva. Az erősen stabilizált köztiterméknek köszönhetően (mélyhőmérsékleten, argon atmoszféra alatt) 90:10 (Z:E) szelektivitás is elérhető a termelés pedig minden esetben elfogadható mértékű (>90%). Az UV-indukált fotoizomerizáció, megfizethető eljárásnak bizonyult a *Z*-izomerek előállítására, azonban az izomer keverék elválasztása speciális felszerelést igényel.

Az előállított vegyületek szerekezetképző sajátságainak vizsgálatára számos műszeres technika kombinációját alkalmaztam, a szilárd kristályos, az oldat fázis és a felületi adszorbeátumok vizsgálatára. *Infravörös spektroszkópia* segítségével felderítettem, hogy mind az *E* mind *Z-fahéjsav* (*3-* és *2,3-arilpropénsavak*) származékok, melyek *fenil*, vagy *öttagú heteroaromás* gyűrűt tartalmaznak, a *karboxil* csoportok közt kialakuló, ciklusos kettős *hidrogénhiddal* összetarott dimereket alkotnak, szilárd fázisban és oldatban , amennyiben az oldószer aprotikus (*kloroform* vagy DMSO). Ettől, cask a *piridil fahéjsavak* mutatnak eltérést, ahol szilárd fázisban sz *ikerionos* szerkezet dominál. Ennek ellenére bebizonyosodott, hogy DMSO-ban lehetséges a *hidrogenhíddal* összetartott dimerek képződése. Az *infravörös spektroszkópiával* nyert eredményeket összehasonlítva (melyeket különböző fázisban rögzítettem) a gyengébb (aromás)C–H…X(aromás)(X = O, S, N, Se), ill. az olefins protonok és a *karboxil* csoport *oxigénje* közt kialakuló másodlagos *hidrogénhidak* jelenléte is valószínűsíthető, melyre az aromás rezgések szisztematikus eltolódásából lehet következtetni.

A telített DMSO oldatban rögzített $^{1}H–^{1}H$ *korrelációs spektrumok* (COSY) azt bizonyították, hogy az előbb említett másodrendű *hidrogénhidak* valóban léteznek, különös tekintettel az aromás gyűrűk közt kialakuló kölcsönhatásra, ahol számos intermolekuláris kölcsönhatást fedeztem fel, nem szomszédos *protonok* közti csatolások révén, melyek átlón kívüli csatolások formájában jelennek meg a spektrumon. Egy, a *hidrogénhidak* jelenlétének további jelét is megfigyeltem: az aromás *protonok* jelszélesedésének formájában, ami sokkal jelentősebb mértékű volt, mint az olefins protonok jelszélesedése. Ez a jelenség azt bizonyítja,a hogy az olefins *protonokkal* kialakuló *hidrogénhidak* szerepe az oldat fázisban nem meghatározó. Az $^{1}H–^{1}H$ *nukleáris Overhauser effektus spektroszkópia*(NOESY) alkalmasnak bizonyult a tanulmányozott vegyületek konformációs állapotainak felderítésére. Az eredmények azt mutatták, hogy a *3-heteroaromás* származékok gyűrűi, híg oldatban (0.05 M in DMSO) szabadon forognak, azonban a *2,3-szubsztituált* változatoknál eltérő volt a tapasztalat. A 2. Szénatomon

található gyűrű viszonyla szabadon foroghat, míg a másik gyűrű konformációs tere valamelyest gátolt.

Az önszerveződő képességet polikristályos fémfelületeken is vizsgáltam (PLD technikával készült *arany* és *ezüst* felületeken) a *dip-coating* eljárást alkalmazva, 0,1 M-os *kloroformos oldatokat alkalmazva*. Az elkészült molekuláris rétegeket az IRM és az AFM technikák kombinációjával tanulmányoztam. Az IRM mérések eredményei igazolták, hogy a molekulák felületen történő megkötése sikeres volt. A adszorbeált rétegekről készült AFM felvételek, számos érdekes, magasabb rendű formációt fedtek fel, ahol a rétegvastagság minden esetben nagyobb volt, mint monomolekulás borítottság esetén. Tehát az említett *hidrogénhidas* kölcsönhatások jelenléte meghatározónak bizonyult. A felületen megtalálható, igen nagy méretű anyagcsúcsok a *π-stacking* kölcsönhatás jelenlétét is valószínűsítik.

A konformáció analízis sikeresen felderítette a *heteroaromás fahéjsav* származékok potenciálfelületét, és az *α-szubsztituált* változatok stabil konformereit. A konformációs terek tulajdonságai teljes egyezést mutattak a NOESY mérések eredményével. Ezt követően a legstabilabb konformereket (melyek a konformáció analízis legkisebb energiájú szerkezetei) magasabb elméleti szinten optimalizáltam, és a *heteroaromás fahéjsavakra* minden esetben teljesen sík szerkezetet kaptam, az *α-szubsztituált* származékokra viszont csavart geometriákat. A kettős *hidrogénhíddal* összetartott dimerekre számolt geometriai adatok és kötési energiák bizonyították, hogy a domináns szerkezeti egység a dimer, átlagosan 15 kcal/mol kötési energiával. Ez az eredmény igen jól korrelál az IR mérések eredményeivel. A molekulák között kialakuló másodrendű kötőerők lehetséges erősségi sorrendjét is sikeresen meghatároztam: C–H…N > C–H…O > C–H…S ~ π–π > C–H…Se, mely igen nagy hasonlóságot mutat a kísérleti eredményekkel.

A felsorolt eredmények egyértelműen azt mutatják, hogy a *heteroaromás fahéjsavak* igen sokoldalú vegyületek. Kiterjedt *hidrogénhidas* rendszereket képeznek minden tanulmányozott fázisban. A már említett felhasználási lehetőségek mellett, a vizsgált származékoknak számos felhasználási lehetősége adódhat az elektronika és a korrózióvédelem területén, köszönhetően a fémfelületeken mutatott viselkedésüknek. Mindemellett számos élettani hatásuk is lehetséges, mely érdekes célponttá teszi őket a gyógyszeripar számára.

ACKNOWLEDGEMENT

Firstly, let me gratefully acknowledge the immense support and help of my supervisor Dr. István Pálinkó, my consultant Dr. Pál Sipos and Dr. Gábor Peintler for the useful suggestions. My grateful thanks go to them for both the theoretical and practical encouragement and teaching I have received from them during the years I have spent in their research group. Let me also gratefully acknowledge the efforts of my colleagues: Éva Bajnóczi, Dr. Mónika Ádok, Eszter Czeglédi, Gábor Varga, Attila Gácsi, Zsolt Ferenc, Dr. Attila Pallagi and Gábor Kozma for their help and suggestions about various problems appearing during my work. Many thanks, to all members of the Material and Solution Structure Research Group. I thank to all who helped me in any aspects in- and outside the University of Szeged, especially for Kai Ruusuvuori from the University of Helsinki for his help with the computational part of this work. Many thanks to my family and friends for the tremendous support they gave me in the last years. Finally, I would like to acknowledge the most important person in my life – my companion, who always supports me, and I especially want to gratefully thank for her patience during the last part of my PhD work.

APPENDIX

Description and preparation of the applied materials:

Reagents:

- **Selenophene** 97% (Sigma-Aldrich, UWFP) **46**
- ***N,N'*-dimethylformamide (DMF)** >99.8% (Merck, stored over 3 Å molecular sieves) **56**
- **Phosphorus oxychloride ($POCl_3$)** 99% (Reagentplus®, Sigma-Aldrich, UWFP)
- **Sodium acetate (NaOAc)** 99% (Sigma-Aldrich, UWFP) **51**
- **Acetic anhydride (Ac_2O)** 95% (RPL, distilled before use) **8**
- **Potassium carbonate (K_2CO_3)** 99% (Reanal, UWFP)
- **Hydrogen chloride (HCl)** 37% (VWR, Ph.Eur.)
- **Sodium hydroxide (NaOH)** (VWR, a.r. grade, used as 10% solution in water)
- **Potassium hydroxide (KOH)** 90% (Molar, reagent grade, UWFP)
- **Sodium in mineral oils** 97% (Merck, rinsed with water-free n-hexane and handled under argon atmosphere, reagent grade)
- **Ethyl acetate (EtOAc)** 99.5% (Fluka, stored over sicc. K_2CO_3)
- **Oxalyc acid ($(COOH)_2 \times 2H_2O$)** 99% (VWR, used in anhydrous form, dried with azeotropic distillation with benzene)
- **Diethyl malonate** 98% (GC) (prepared from anhydrous oxalic acid via Fischer esterification with EtOH)**19**
- **Piperidine** 99% (Reanal, stored over KOH flakes) **20**
- **Tetrabutyl ammonium bromide (TBAB):** 98% (BDH, UWFP) **53**
- **Pyridine** 99% (Sigma-Aldrich, ACS reagent®, stored over NaOH pellets) **43**
- **1,4-hydroquinone** 99% (Merck, UWFP) **54**
- **Malonic acid** 99% (VWR, UWFP) **22**
- Each aromatic aldehyde was vacuum-distilled before use.
- **Benzaldehyde** 99% (Reagentplus®, Sigma-Aldrich) **7**
- **2-furancarboxaldehyde (furfural)** 99% (ACS reagent®, Sigma-Aldrich) **85**
- **3-furanecarboxaldehyde** 97% (Sigma-Aldrich) **86**
- **2-pyridinecarboxaldehyde** 99% (Sigma-Aldrich) **87**
- **3-pyridinecarboxaldehyde** 98% (Sigma-Aldrich) **88**

- **4-pyridinecarboxaldehyde** 97% (Sigma-Aldrich) **89**
- **2-thiophenecarboxaldehyde** 98% (Sigma-Aldrich) **90**
- **3-thiophenecarboxaldehyde** 98% (Sigma-Aldrich) **91**
- **1-Octanol** 99% (Sigma-Aldrich, reagent grade, UWFP) **57**
- **4-dimethylaminopyridine (DMAP)** 99% (Sigma-Aldrich, UWFP) **59**
- **Triethylamine (TEA)** 99% (Sigma-Aldrich, stored over NaOH pellets)
- **Bromine (Br_2)** 99% (Sigma-Aldrich, used as 10 M solution in dichloromethane)
- **Sulfuric acid (H_2SO_4)** 98% (VWR, UWFP)
- **Palladium chloride ($PdCl_2$)** 99% (Reanal, at., UWFP)
- **Lead acetate ($Pb(OAc)_2$)** 99% (Reanal, used as 7% aqueous solution)
- **Formic acid (HCOOH)** 99% (Fluka, freshly distilled before use)
- **Calcium chloride ($CaCl_2$)** 96% (Reanal, anhydrous, granulated, UWFP)
- **Sodium carbonate ($NaCO_3$)** 99% (Fluka, used as 20% aqueous solution)
- **Barium hydroxide ($Ba(OH)_2 \times 8H_2O$)** 98% (Reanal, used as 0.8M aqueous solution)
- **Calcium carbonate ($CaCO_3$)** purity was not checked, (prepared via the precipitation of $CaCl_2$ with Na_2CO_3).
- **Barium sulfate ($BaSO_4$)** purity was not checked, (prepared via the precipitation of $Ba(OH)_2$ with H_2SO_4).
- **Hydrogen (H_2)** 99.98% (Messer, in 6m^3, 150bar gas cylinder, UWFP)
- **Quinoline (2,3-benzopyridine)** 98% (Reachim, used after vacuum distillation) **27**
- **3,6-dithia-1,8-octanediol** 98% (Sigma-Aldrich, UWFP) **28**
- **Chloroacetic acid** 97% (Reachim, UWFP)
- **Methyl chloroacetate** 99% (GC) (prepared via Fischer esterification of chloroacetic acid with MeOH) **62**
- **Sodium Iodide (NaI)** 99% (Reanal, anhydrous, dried at 120°C)
- **Trimethyl phosphite ($P(OMe)_3$)** 97% (Sigma-Aldrich, UWFP) **65**
- **Iodine** 99.5% (Merck, reagent grade, UWFP)
- **Sodium thiosulfate ($Na_2S_2O_3 \times 5H_2O$)** 98% (VWR, UWFP)
- **Trimethylsilyl chloride (TMSCl)** 99% (Sigma-Aldrich, UWFP) **66**

- **Oxalyl chloride ($(COCl)_2$)** 95% (GC) (prepared via the halogenation of anhydrous oxalic acid with PCl_5) **68**
- **Phosphorus pentachloride (PCl_5)** 95% (Fluka, UWFP)
- **2,2,2-Trifluoroethanol** 99% (Sigma-Aldrich, Reagentplus®, UWFP) **70**
- **O-cresol** 99% (Sigma-Aldrich, Reagentplus®, UWFP) **71**
- **Potassium in mineral oils** 95% (Merck, rinsed with water-free hexane and handled under argon atmosphere, reagent grade)
- ***t*-Butanol** 99% (Reanal, stored over sicc. $MgSO_4$)
- **Potassium t-butoxide (*t*-BuOK)** purity was not checked, (prepared via the reaction of potassium with t-butanol) **72**
- **Potassium bis-trimethylsilylamide (KHMDS)** 95% (Sigma-Aldrich, stored under argon atmosphere) **32**
- **18-crown-6 (complexone)** 99% (Sigma-Aldrich, UWFP) **31**
- **Sodium hydride (NaH)** 95% (Sigma-Aldrich, dry, stored under argon atmosphere at −20°C)
- **Phenylacetic acid** 99% (Sigma-Aldrich, UWFP) **17**
- **Potassium cyanide** 96% (Reanal, UWFP)
- **Red phosphorous** 97% (Sigma-Aldrich, UWFP)
- **Phosphorus tribromide*** 98% (GC) (prepared via the halogenation of red phosphorous with bromine)

Supporting materials:

- **Sodium hydrogencarbonate ($NaHCO_3$)** 99.7% (Sigma-Aldrich, Reagentplus®, used as 20% aqueous solution)
- **Sodium chloride (NaCl)** 99.9% (Reanal, a.p., used as saturated aqueous solution)
- **Ammonium chloride (NH_4Cl)** 99% (Reanal, used as 10% aqueous solution)
- **Sodium sulfate (Na_2SO_4)** 99% (Fluka, anhydrous, UWFP)
- **Magnesium sulfate ($MgSO_4$)** 98% (Reanal, anhydrous, UWFP)
- **Benzophenone** 99% (Fluka, UWFP)
- **3 Å molecular sieves** (Sigma-Aldrich, 1.6mm beads, regenerated at 300°C)
- **4 Å molecular sieves** (Fluka, rods, regenerated at 300°C)
- **Silica gel: Normasil 60 (43- 63µm)**(VWR, for column chromatography)

- **Silicagel on TLC Al foils 60 Å** (Sigma-Aldrich, with fluorescent indicator)

Solvents:

- **Methanol (MeOH)** 99.9% (Sigma-Aldrich, Reagentplus®, UWFP)
- **Ethanol (EtOH)** 99.8% (VWR, anhydrous, stored over 3 Å molecular sieves)
- **Tetrahydrofuran (THF)** 99.9% (VWR, anhydrous, stored over Na/benzophenone, distilled from NaH before use)
- **Ethyl acetate (EtOAc)** 99.9% (Sigma-Aldrich, Chromasolv®, HPLC grade) **51**
- **Acetone** 99% (VWR, UWFP) **63**
- **Chloroform ($CHCl_3$)** 99% (Sigma-Aldrich, contains 1% ethanol as stabilizer)
- **Dichloromethane (CH_2Cl_2)** 99.5% (Sigma-Aldrich, contains 0.5% amylene as stabilizer, stored over sicc. $CaCl_2$)
- **Chlorobenzene** 99% (Fluka, stored over sicc. $CaCl_2$) **55**
- **n-hexane** 96% (Sigma-Aldrich, anhydrous, UWFP)
- **Benzene** 99% (Merck, stored over sodium wire)
- **o-Xylene** 99% (Merck, puriss, stored over sodium wire) **52**
- **Diethyl ether** 99% (VWR, stored over sodium wire)
- **DMSO (dimethyl sulfoxide)** 99.8% (Merck, spectroscopic grade)

The C source code of the driving unit of the UV-flow reactor:

```
#define F_CPU 1000000UL
#include <avr/io.h>
#include <util/delay.h>
int main(void)
{
    DDRD |= (1 << PD5);                 // define the "Gate" (G) of the FET is on PD5 pin
    while(1)
    {
        if (ACSR & (1 << ACO)) {      // comparator output high? switch on if comparator output is high
            PORTD |=  (1 << PD5);       // switch PD5 (pump) on trough the FET
             delay ms(2000);           // hold for 2 seconds (change the solution in the cuvette)

            PORTD &= ~(1 << PD5);       // switch PD5 off (stop the pump)
            _delay_ms(21600000); // hold for 6 hours (UV irradiation period can be varied)
        }
        else {
                                        // comparator output low, so switch PD5 off
            PORTD &= ~(1 << PD5);       // switch PD5 (pump) off trough the FET
        }
    }
}
```

The simplified standalone (AutoHotKey) source code of Quantum Hyppo (demo), without the GUI section:

```
SetBatchLines, -1                    ;//the program will never sleep
#Noenv                                ;//avoids checking empty variables
SetWorkingDir %A_ScriptDir%           ;//ensures the consistent starting directory
#SingleInstance Off          ;// allows multiple instances of the script to run concurrently
#Include tf.ahk                       ;//code library for ASCII files v.3.4

FileSelectFile, input, 1, , Please choose a Gaussian input file., Gaussian input files (*.gjf)
FileSelectFolder, inputfolder, , ,Please select a folder for new Gaussian input files.
FileSelectFolder, outputfolder, , ,Please select a folder for new Gaussian output files.
FileSelectFile, batch, S, , Choose a name for the batch file!, (*.bcf)
FileCreateDir, %A_ScriptDir%\Temp  ;//create a directory for temporary files
                                      ;//multiple selection windows to determine files
SplitPath, batch,,batchfolder,, batchname   ;//split the file name from the file route
dash=\                                ;//dash cannot be used directly in some cases

FileAppend,                           ;//creating the batch file and it's structure
(
!
!%batchname%
!start=1

), %batchfolder%%dash%%batchname%.bcf
                                      ;//user can defined these variables (trough GUI)
min= -180.0                           ;//lower dihedral angle value
max= 180.0                            ;//higher dihedral angle value
vari= 128                             ;//simultaneous variations on the dihedral angles

FileRead, original, %input%           ;//read in file content
Loop, read, %input%,                  ;//read the input line by line
       {
       col1:=TF_Colget(original, A_Index, A_Index, 1, 1) ;//get the first column
       Ifinstring, col1, D, fileappend, %A_Index%`n, %A_Scriptdir%\Temp\dnum.txt
       }
                                      ;//find the dihedral angle section and store it
FileRead, dlines, %A Scriptdir%\Temp\dnum.txt
dihed:=TF CountLines(dlines)          ;//count the dihedral sets
dihed:=dihed-1
FileReadLine, endline, %A_Scriptdir%\Temp\dnum.txt, 1
endline:= endline - 1                 ;//determine the last line
section1:=TF Colget(original, 1, endline, 1, 100000)
                                      ;//get all other lines from the input file
Loop, read, %A_scriptdir%\Temp\dnum.txt
       {
       FileReadLine, dihedline, %input%, %A Loopreadline% ;//get the dihedral set from input
                     lenght:=StrLen(dihedline)     ;//get the length of the line
                     atomnum:=TF Colget(dihedline,1,1,3,lenght,0) ;//get the dihedral columns
                     Stringsplit, atom, atomnum, %A Space%,
                     A1:=atom1        ;//split the line to 4 separate atom numbers
                     A2:=atom2
                     A3:=atom3
                     A4:=atom4
                     Fileappend,%A1%%A Space%%A2%%A Space%%A3%%A Space%%A4%`n
,%A_scriptdir%\Temp\atoms.txt
       }                              ;//create a file with the atomic numbers

Splitpath, input, , , , namesort      ;//determine the file name for the new files
Loop, %vari%
       {
       SetFormat, float, 05.0         ;//set the floating point format to 00000
       name += 1.0                    ;//add 1 to the file name in each cycle
       namelong = %name%_%namesort%   ;//final name format: 00001originalfilename.gjf
       FileDelete, %A scriptdir%\Temp\dihedrals.txt
       Loop, %dihed%
              {
              Filereadline, atoms, %A_Scriptdir%\Temp\atoms.txt, %A_Index%
              If atoms=
                     Break
              Else
              rand := 0                ;//empty the variable
              SetFormat, float, 0.6
              Random, rand, %min%, %max%    ;//generate random number with 0.000000 format
              Fileappend, D %atoms% %rand% F`n, %A_scriptdir%\Temp\dihedrals.txt
              }                        ;//create the new dihedral section of the input
```

```
Fileread, dihedrals, %A_scriptdir%\Temp\dihedrals.txt
Fileappend, %section1%`n%dihedrals%, %inputfolder%\%namelong%.gjf
                                    ;//merge the file parts and create the input new file
extin=.gjf
extout=.out
infile:=inputfolder dash namelong extin
outfile:=outputfolder dash namelong extout
separator= ,                  ;//add the route of the input and output files to the batch file
Fileappend, %infile%%separator% %outfile%`n, %batchfolder%%dash%%batchname%.bcf
       }
FileDelete, %A_scriptdir%\Temp\dihedrals.txt              ;//delete temporary files
FileDelete, %A_Scriptdir%\Temp\atoms.txt
FileDelete, %A_Scriptdir%\Temp\dnum.txt
Msgbox, 4160, Finished!, All desired files were created.   ;//finish message
Return                                                     ;//close the program
```

The simplified standalone (AutoHotKey) source code of Quantum Hyppo (demo), without the GUI section:

```
SetBatchLines, -1                          ;//the program will never sleep
#Noenv                                     ;//avoids checking empty variables
SetWorkingDir %A_ScriptDir%                ;//ensures the consistent starting directory
#SingleInstance Off                        ;// allows multiple instances of the script to run
concurrently
#include tf.ahk                            ;//code library for ASCII files v.3.4

kcal = 627.509469                          ;//define  the  conversion  between  Hartree  and
kcal/mol units
                                           ;//select the folder that contains output files
FileSelectFolder, outfolder, 1, Select folder, Please select a folder that contains Gaussian
output files!
FileCreateDir, %A_ScriptDir%\Temp          ;//create a directory for temporary files

Loop, %outfolder%\*.out                    ;// read output files one by one
       {
       actpath:=A_Loopfilefullpath
       Loop, read, %actpath%               ;//read the files line by line
              {                            ;//searching for errors
              Ifinstring, A Loopreadline, Error termination
                     {                     ;//if  an  error  founded  copy  the  file  to  Error
directory
                     FileCreateDir, %outfolder%\Error
                     FileCopy, %A_Loopfilefullpath%,  %outfolder%\Error
                     FileAppend, %A Loopfilefullpath%`n, %outfolder%\Error\errorlist.txt
                     }                     ;//create  a  list  with  the  names  of  the  erroneous
files
              Ifinstring, A_Loopreadline, Erroneous write
                     {                     ;//search for other problems
                     Filecreatedir, %outfolder%\Error
                     Filecopy, %A Loopfilefullpath%,  %outfolder%\Error
                     Fileappend, %A Loopfilefullpath%`n, %outfolder%\Error\errorlist.txt
                     }                     ;//search for other problems
              Ifinstring, A_Loopreadline, could not allocate memory
                     {
                     Filecreatedir, %outfolder%\Error
                     Filecopy, %A Loopfilefullpath%,  %outfolder%\Error
                     Fileappend, %A Loopfilefullpath%`n, %outfolder%\Error\errorlist.txt
                     }
              }
       }                                   ;//read in the first line of error list
FileReadLine, errorline, %outfolder%\Error\errorlist.txt, 1
If errorline=                              ;//if no errors found, send check message
       MsgBox, 4160, Checked!, No output files terminated with error.
Else                                       ;//if errors found, send distress message
       MsgBox, 4144, Checked!, Output files terminated with error moved to Error directory.
Loop, %outfolder%\*.out                    ;//read output files
       {
       Splitpath, A Loopfilefullpath, , , ,actfilename     ;//get the actual file name
       exgeomfile:= A_Loopfilefullpath
       Loop, read, %exgeomfile%            ;//read the output file content
              {
              Ifinstring, A Loopreadline, Input Orientation
                     {                     ;//find the optimized geometry section
                     startline:= A_Index+5
```

```
                        }
                }
        Loop
                {
                Filereadline, geomline, %exgeomfile%, %startline%
                startline +=1                   ;//identify the end of the geometry section
                Ifinstring, geomline, -----------------------
                        {
                        lines := A_Index-1 ;//write the number of the atoms to the geometry file
                        Fileappend,   %lines%`n   %actfilename%`n,   %outfolder%\%actfilename%.xyz

                        startline := startline-A_Index
                        Break
                        }
                }
        Loop
                {
                x =                             ;//empty the variable of coordinates
                y =
                Filereadline, geomline, %exgeomfile%, %startline%
                startline +=1                   ;//read the geometry lines one by one
                Ifinstring, geomline, -----------------------
                        Break
                Else
                        {
                        needpart =              ;//split the line to components
                        stringsplit, part, geomline, %A_Space%, %A_Space%%A_Tab%
                        Loop, %part0%
                                {
                                actnumberall:= part%A_Index%
                                If actnumberall <>      ;//check the components
                                        {
                                        needpart += 1
                                        }
                                If needpart = 2         ;//the first geometry data line
                                        {
                                        actnumber= %actnumberall%
                                        If actnumber = 1
                                                atomsymbol:= "H"
                                        else if actnumber = 6
                                                atomsymbol:= "C"
                                        else if actnumber = 7
                                                atomsymbol:= "N"
                                        else if actnumber = 8
                                                atomsymbol:= "O"
                                        else if actnumber = 16
                                                atomsymbol:= "S"
                                        else if actnumber = 34
                                                atomsymbol:= "Se"
                                        }       ;//translate the atom numbers to atom symbols
                                If needpart = 4
                                        {
                                        if x <>
                                        x = %x%         ;//get x-coordinate
                                        Else
                                        x = %actnumberall%
                                        }
                                If needpart = 5
                                        {
                                        if y <>
                                        y = %y%         ;//get y-coordinate
                                        Else
                                        y = %actnumberall%
                                        }
                                If needpart = 6
                                        {
                                        z = %actnumberall%      ;//get z-coordinate
        Fileappend, %atomsymbol%%A_Tab%%x%%A_Tab%%y%%A_Tab%%z%`n, %outfolder%\%actfilename%.xyz
                                        break
                                        }       ;//create the geometry file line byline
                                }
                        }
                }
        }                                       ;//create the header of the results file
Fileappend,  File  name%A_tab%Corrected  ZPE%A_tab%Lowest  Frequency%A_Tab%1.  dihedral%A_Tab%2.
dihedral`n, %outfolder%\results.txt
Loop, %outfolder%\*.out                         ;//read again the output files
```

```
	{
	actindex:= A Index
	actpath:= A Loopfilefullpath
	Splitpath, A Loopfilefullpath, , , ,actfilename     ;//get the actual file name
	Loop, read, %actpath%                        ;//read the actual file line by line
		{
		Ifinstring, A Loopreadline, Zero-point correction=
			{                                ;//find the thermal correction for energy
		StringReplace, sortedzecorr, A Loopreadline, Zero-point correction=,, all
		Stringreplace, sortedzecorr, sortedzecorr, (Hartree/Particle),, all
		Stringreplace, sortedzecorr, sortedzecorr,%A_Space%,,all
			}
		Ifinstring, A Loopreadline, Sum of electronic and zero-point Energies=
			{                                ;//find the zero-point energy
StringReplace, sortedze, A Loopreadline, Sum of electronic and zero-point Energies=,, all
			Stringreplace, sortedze, sortedze,%A Space%,,all
			ZPE:=((sortedze - sortedzecorr) * kcal )
			}                        ;//calculate the corrected energy in kcal/mol
		Ifinstring, A Loopreadline, and normal coordinates:
			{                                ;//search for frequencies
			frfound =
			B_Index:=A_Index
			Loop,
				{
				if frfound =1
					Break
				Else
					Filereadline, frline, %actpath%, %B_Index%
				Ifinstring, frline, Frequencies --
					{                ;//find frequencies
					frfound = 1      ;//split the frequency line to components
					StringReplace, freqall, frline, Frequencies --,, all
					stringsplit, freqlow, freqall, %A_Space%, %A_Space%
					Loop, %freqlow0%
						{
						fract:= freqlow%A Index%
						if fract <>
							{ ;//floor the frequency to whole number
							FRACT:= Floor(fract)
							break
							}
						}
					}
				Else B_Index +=1
				}
			}
	Ifinstring, A_loopreadline, The following ModRedundant input section has been read:
			{                                ;//search for the original dihedral set
			rownumber:= A index
			add= 0
			Loop,
				{
				add +=1
				actrow:=(rownumber+add)
				Filereadline, dline, %actpath%, %actrow%
				ifinstring, dline, D          ;//find dihedral sets
					{
					Stringreplace, dline, dline,D,,all
					Stringreplace, dline, dline,F,,all
					Loop, parse, dline, %A Space%
						{                ;//trim and split the dihedral line
						If A_Loopfield <>
							{
			Fileappend, %A Loopfield%%A Space%, %A scriptdir%\Temp\trimmed.txt
							}
						}
					Filereadline, trimmed, %A_scriptdir%\Temp\trimmed.txt, 1
			Stringsplit, a, trimmed, %A_Space%,. ;//write out the atomic numbers
Fileappend, %a1%%A Space%%a2%%A Space%%a3%%A space%%a4%`n, %A scriptdir%\Temp\datoms.txt
					Filedelete, %A scriptdir%\Temp\trimmed.txt
					Fileappend, %a5%`n, %A scriptdir%\Temp\dihedrals.txt
					}                        ;//write out the dihedral angle
				Else
					Break
				}
			}
		}
```

```
        FileReadLine, d1, %A_scriptdir%\Temp\datoms.txt, 1  ;//read in dihedral components
        FileReadLine, d2, %A scriptdir%\Temp\datoms.txt, 2
        FileReadLine, a1, %A scriptdir%\Temp\dihedrals.txt, 1
        FileReadLine, a2, %A scriptdir%\Temp\dihedrals.txt, 2
Fileappend,
%actfilename%%A_tab%%ZPE%%A_tab%%FRACT%%A_Tab%%d1%%A_Tab%%a1%%A_Tab%%d2%%A_Tab%%a2%`n,
%outfolder%\results.txt
        FileDelete, %A scriptdir%\Temp\datoms.txt ;//create the results file with several data
        FileDelete, %A scriptdir%\Temp\dihedrals.txt
        }
MsgBox, 4160, Finished!, All desired data were extracted. ;//finish message
Return                                                    ;//close the program
```