

# UNIVERZITA PALACKÉHO V OLMOUCI

Přírodovědecká fakulta

Katedra analytické chemie



**Vícerozměrné separační techniky pro analýzu metabolomu lidské moči**

## DISERTAČNÍ PRÁCE

Autor práce:

RNDr. Jaroslava Jáčová

Vedoucí práce:

prof. RNDr. Tomáš Adam, Ph.D.

Studijní obor:

Analytická chemie

Forma studia:

Prezenční

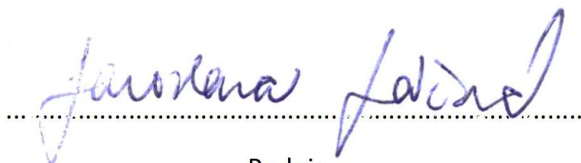
Olomouc 2019

**Prohlášení:**

Prohlašuji, že jsem závěrečnou práci zpracovala samostatně a že jsem uvedla všechny použité informační zdroje a literaturu. Tato práce ani její podstatná část nebyla předložena k získání jiného nebo stejného akademického titulu.

Souhlasím s tím, aby má práce byla zpřístupněna v knihovně Katedry analytické chemie Přírodovědecké fakulty Univerzity Palackého v Olomouci a v informačním systému Univerzity Palackého v Olomouci.

V Olomouci 21. 6. 2019



Podpis

### **Poděkování:**

Ráda bych poděkovala prof. Tomáši Adamovi za to, že mne přijal do svého týmu, Jean-Marie D. Dimandjovi, za uvedení do vědecké komunity GC×GC, Davidu Friedeckému, že se mne ujal, Petrovi Wojtowiczovi, za všechna vodítka, která mi zanechal, Alžbětě Gardlo za generování modelů, pomoc s matematikou a R-softwarem, Vojtovi Bekárkovi za seznámení s plynovou chromatografií, Jitce Zrostlíkové, Tomáši Kovalczukovi, Markovi Merrickovi a Jackovi Cochranovi za cenné rady, Hance Janečkové, Rádi Karlíkové, Vladce Růžičkové, Katce Mičové a Evě Hlídkové za podporu, a všem, co mne mají rádi, za to, že jsou...

Disertační práce vznikla za finanční podpory projektů GAČR 15-34613L, NPU I (LO1304) a GAČR 18-12204S.

## OBSAH

1	Metabolomika .....	9
2	Metabolom moči .....	9
3	Metabolomický experiment .....	10
4	Instrumentální analytické techniky v metabolomice.....	12
4.1	Výpočet ortogonalita dvojrozměrných chromatografických systémů .....	13
4.2	Škálování separačního prostoru.....	17
4.3	Dimenzionalita vícerozměrných separací.....	18
5	Sběr a příprava vzorků .....	20
6	Vyhodnocování metabolomických dat.....	22
7	Cíle práce.....	28
8	Imobilizovaná ureasa pro GC-MS metabolomiku moči .....	29
8.1	Metody přípravy vzorku pro analýzu metabolomu moči pomocí GC-MS .....	29
8.2	Materiál .....	30
8.3	Vzorky.....	30
8.3.1	Příprava vzorků COA pro GC-MS analýzu .....	32
8.3.2	Příprava vzorků COA pro LC-MS/MS analýzu.....	33
8.3.3	Příprava vzorků pacientů pro metabolomický experiment .....	33
8.4	Metody .....	34
8.4.1	GC-MS analýza a procesování dat.....	34
8.4.2	LC-MS analýza a procesování dat.....	35
8.5	Výsledky a diskuse.....	36
8.5.1	Pracnost metod.....	36
8.5.2	Kvalitativní změny metabolomu .....	36
8.5.3	Artefakty.....	43
8.5.4	Kvantitativní změny metabolomu.....	47
8.5.5	Analytická variabilita .....	53
8.5.6	Reprodukovatelnost retenčních časů.....	55
8.5.7	Metabolomický experiment s močemi pacientů s glutarovou acidurií I .....	55
8.6	Závěr.....	57
9	WOSEL jako nový způsob škálování 2D separačního prostoru.....	58
9.1	Způsoby škálování separačního prostoru .....	58
9.2	Metody .....	60

9.2.1	Syntetické chromatogramy.....	60
9.2.2	Reálné chromatogramy .....	60
9.2.3	Výpočet ortogonalita.....	61
9.3	Výsledky a diskuse.....	61
9.3.1	Volba syntetických modelů.....	61
9.3.2	Vliv metody škálování 2D prostoru na výslednou ortogonalitu.....	62
9.3.3	Vliv vzorku použitého pro optimalizaci na škálování .....	66
9.3.4	Vliv způsobu škálování na různé přístupy výpočtu ortogonalita .....	67
9.3.5	Specifika vymezení separačního prostoru pomocí WOSEL .....	69
9.4	Závěr .....	69
10	Význam dimenzionalita vzorku pro výpočet ortogonalita 2D separací.....	70
10.1	Lokální a globální ortogonalita.....	70
10.2	Metody .....	71
10.2.1	Syntetické chromatogramy.....	71
10.2.2	Reálné chromatogramy .....	71
10.2.3	Výpočet ortogonalita.....	72
10.3	Výsledky a diskuse.....	72
10.3.1	Rozsah deskriptorů.....	73
10.3.2	Nízkodimenzionální vzorky .....	77
10.3.3	Vysokodimenzionální vzorky.....	83
10.3.4	Požadovaná dimenzionalita .....	86
10.4	Závěr .....	88
11	Nalezení nejvíce ortogonální kombinace kolon pro GC×GC analýzu metabolomu moči.....	89
11.1	Kolony v GC×GC.....	90
11.2	Chemikálie .....	91
11.3	Příprava vzorku .....	93
11.3.1	Moč.....	93
11.3.2	Plasma .....	93
11.4	Metody .....	93
11.5	Výsledky a diskuse.....	94
11.5.1	Určení ortogonálního kolonového uspořádání pro GC×GC analýzu moči .....	94
11.5.2	Určení ortogonálního kolonového uspořádání pro GC×GC analýzu plasmy ....	102
11.5.3	Zamyšlení nad ortogonalitou v optimalizaci GC×GC systémů .....	106
11.6	Závěr .....	107
12	Závěr.....	108

13	Výhledy.....	109
14	Seznam použité literatury .....	110
15	Seznam zkratk.....	126
16	Doplňující informace .....	128
17	Přílohy .....	147

## Souhrn

Předkládaná disertační práce se zabývá instrumentálními analytickými metodami v metabolomice. Metabolomika je vědní disciplína umožňující detailní pohled na celkový profil nízkomolekulárních analytů (metabolitů) v biologických materiálech. Analytické nástroje proto musí čelit velké rozmanitosti a koncentračnímu rozsahu sledovaných analytů při vysoké komplexitě vzorků.

V první části je vyvinut nový přístup přípravy vzorku na základě imobilizované ureasy, který umožňuje maximální zachování metabolomu moči při eliminaci obvyklých artefaktů a interferentů. Nově vyvinutá metoda je porovnána s relevantními již dříve publikovanými postupy (volnou ureasou, kyselou extrakcí do ethylacetátu) a vzorky ponechanými bez úpravy. Přístup založený na imobilizované urease je úspěšně aplikován na necílenou metabolomickou analýzu vzorků moči pacientů s dědičnou metabolickou poruchou - glutarovou acidurií typu I - a zdravých kontrol.

V rámci zaměření na vícerozměrné separační metody, zejména GC×GC, je představen nový způsob škálování dvojdimenzionálního separačního prostoru, WOSEL (z angl. whole separation space scaling) založený na analytické metodě. Při aplikaci WOSEL je separační prostor ohraničen mrtvým časem na jedné straně a koncem separačního záznamu na straně druhé. Oproti tomu u tradičních přístupů je separační prostor zcela nebo z části definován retenčním časem analytů (nejvíce a nejméně zadržovaným analytem daného vzorku). Oba dříve publikované přístupy jsou proto závislé na vzorku použitém k optimalizaci. Při škálování pomocí WOSEL lze díky tomu s výhodou použít jakýkoli vzorek podobného typu. Navíc nedochází ke zkreslení separačního prostoru (i po škálování odpovídá vzhledu chromatogramu).

Dále je pozornost věnována nalezení optimálního způsobu určení ortogonality dvojdimenzionálních chromatografických systémů. Je testováno celkem osm deskriptorů, které ortogonalitu počítají na základě odlišných principů. Žádný z deskriptorů ale není schopen ohodnotit homogenitu pokrytí (lokální ortogonalitu) a rozsah využitého separačního prostoru (globální ortogonalitu) najednou. Proto je pomocí geometrického průměru zkombinován nejlepší deskriptor lokální a globální ortogonality do nového deskriptoru nazývaného ASCA (z angl. AN and SC average), zahrnující oba aspekty ortogonality. Pomocí tohoto deskriptoru je nalezena optimální kombinace kolon pro GC×GC-MS analýzu metabolomu moči a plasmy.

## Summary

The thesis deals with instrumental analytical methods in metabolomics. Metabolomics is a scientific discipline that focuses on the low molecular weight analytes (metabolites) contained in biological materials. Therefore, analytical tools must face the great diversity and concentration range of the analytes analyzed and a high sample complexity.

In the first part of the thesis, a new approach to sample preparation based on immobilized urease is developed, which allows maximum retention of urine metabolome by eliminating common artefacts. This newly developed method is compared with methods already published (free urease, acidic extraction into ethyl acetate) and samples left untreated. The approach based on immobilized urease has been successfully applied to untargeted metabolomic analysis of urine samples from patients with inherited metabolic disorder – glutaric aciduria type I - and healthy controls.

Focused on multidimensional separation methods, especially GC × GC, a new method of scaling of a two-dimensional separation space, Whole Separation Space Scaling (WOSEL), is introduced. When using WOSEL, the separation space in both dimensions begins with the void time and ends with the last time point of data acquisition. Compared to WOSEL traditional scaling methods applies separation space delineated by the retention time of the most (and in one case least) retained sample analyte. Both previously published approaches are dependent on the sample used for optimization. In the case of WOSEL, any sample of a given (or similar) type of biological material can be used. Moreover, the space in which orthogonality is calculated corresponds to the perceived quality of the chromatogram (no separation space distortion occurs).

Furthermore, attention is paid to finding the optimal way of determining orthogonality of two-dimensional chromatographic systems. A total of eight descriptors are tested that calculate orthogonality based on different principles. However, none of the descriptors is able to evaluate homogeneity of coverage (local orthogonality) and the extent of separation space used (global orthogonality) at once. Therefore, using the geometric mean, the best local and global orthogonality descriptor is combined into a new descriptor called ASCA (i.e. AN and SC average), encompassing both aspects of orthogonality. Using this descriptor, an optimal column combination for GC × GC-MS urine and plasma metabolome analysis is evaluated.



## 1 Metabolomika

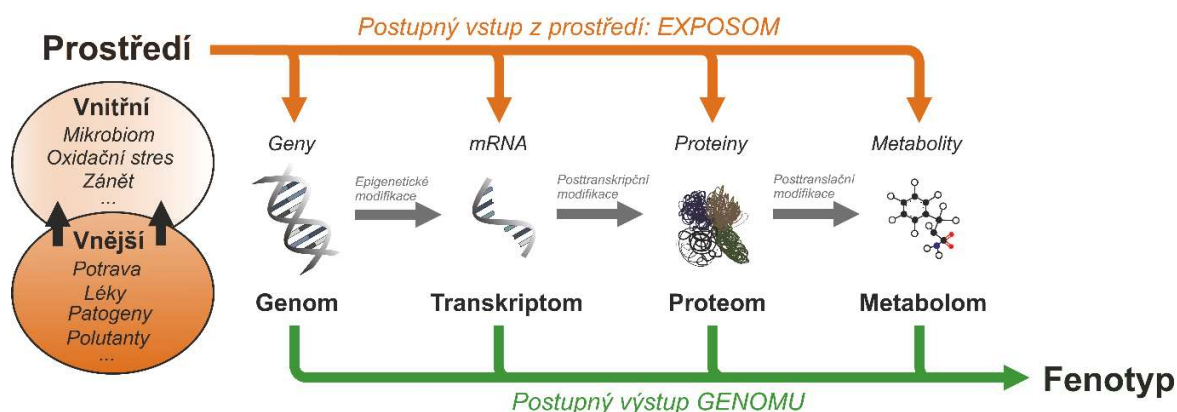
Studiem metabolomu se zabývá věda nazývaná metabolomika. Prostřednictvím specifických metabolických profilů studovaných systémů nabízí lepší porozumění jejich biochemickým a patobiochemickým procesům [1], odhaluje mechanismy onemocnění [2], identifikuje nové biomarkery [3] a umožňuje personalizovanou léčbu onemocnění [4] či sledování účinků terapie [2,5].

Metabolomika jako vědní disciplína hledá rozdíly ve složení komplexního profilu nízkomolekulárních látek mezi alespoň dvěma skupinami vzorků (nejčastěji mezi skupinou pacientů a zdravých kontrol). Řeší dva základní typy problémů: buď sleduje a porovnává abundance předem vybraných, známých metabolitů u sledovaných metabolických drah, tzv. cílená metabolomika, nebo ve vzorcích vyhledává a zaznamenává všechny látky vybraného rozsahu molekulových hmotností bez prvotní znalosti jejich identifikace, tzv. necílená metabolomika [6].

Cílená metabolomika tak přináší informace o využití substrátů a produktů metabolických reakcí, fungování metabolických drah sledovaného biologického systému, umožňuje popsat a hlouběji porozumět jeho biochemii a patobiochemii [7]. Oproti tomu necílená metabolomika vzhledem k obvykle vyššímu pokrytí metabolomu umožňuje formulovat nové hypotézy, odhalit nové biochemické dráhy a biomarkery onemocnění [8]. Oba přístupy poskytují z části komplementární informace, a proto se často provádějí současně.

## 2 Metabolom moči

Metabolom (Obr. 1) je soubor všech malých molekul [9] s molekulovou hmotností do 1500 Da [10] přítomných v daném okamžiku v biologickém materiálu (organismu, orgánu, tkáni, buňce, mozkomíšním moku, krevní plasmě či séru, moči, stolici, potu, slzách apod.). Poskytuje důležitou informaci o aktuálním stavu biologického systému (jeho fenotypu), protože je odrazem jeho genetického vybavení (genomu), aktuálního stavu (transkriptomu a proteomu), a nakonec také vlivu prostředí (exposomu) [11].



Obr. 1: Metabolom jako nejvěrnější obraz fenotypu na molekulární úrovni (převzato z [11])

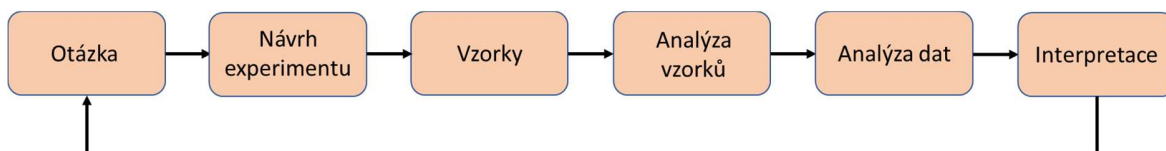
Metabolom je heterogenní skupinou chemických látek, jejichž společným znakem je pouze nízká molekulová hmotnost a výskyt v daném biologickém materiálu. Např. metabolom člověka obsahuje látky 356 různých chemických tříd [12]. V moči se vyskytují především látky hydrofilní, zatímco v krvi převažují lipidy a mastné kyseliny, které v ledvinách do moči téměř nepřecházejí. Zastoupení hydrofilních látek v moči a séru je téměř identické, přesto lze dle Bouarty a kol. [13] v moči detekovat přibližně o 500 látek více než v krvi. Příčinou je koncentrační schopnost ledvin, které při zpětném vstřebávání vody nepřímo zvýší koncentraci metabolitů v moči až o tři řády [13,14]. Navíc v moči nefungují homeostatické mechanismy, proto mohou být rozdíly v hladinách metabolitů mnohem výraznější a lépe tak vypovídají o klinickém stavu organismu [15,16]. To, společně se snadným (neinvazivním) odběrem, činí moč vyhledávaným diagnostickým materiálem [14,15,17].

### 3 Metabolomický experiment

V metabolomických experimentech se sleduje zastoupení mnoha různých analytů ve vzorcích, které jsou v případě humánních vzorků zatíženy vysokou interindividuální variabilitou (vliv patofyziologického stavu jedince, výživy, času odběru, pohlaví, věku, medikace, aj.). Všechny faktory je nutné zvážit a experiment dobře naplánovat. Správné plánování experimentu je nejdůležitější částí metabolomické analýzy [8]. Udělá-li se chyba hned na začátku experimentu, její důsledky jsou fatální, protože následky malého primárního omylu se v každém dalším kroku zesilují, až skončí zkreslenými nebo úplně nesprávnými výsledky [18]. Typické podcenění designu experimentu obvykle vede k nízké síle studie.

Každý metabolický experiment (Obr. 2) musí začít správně položenou otázkou: „Co chceme zjistit?“ Podle ní se po podrobném studiu literatury (které nám rozkryje, co už se ví) zvolí biologický

materiál, ve kterém se bude měřit obsah metabolitů. Podle spektra či typu sledovaných metabolitů pak metabolomický přístup, způsob přípravy vzorku, analytický systém, metoda analýzy i zpracování dat [19]. Na konci experimentu by interpretace výsledků měla přinést odpověď na úvodní otázku.



Obr. 2: Schéma metabolomického experimentu (převzato z [19]).

Vlastní provedení experimentu se pro jednotlivé metabolomické přístupy výrazně odlišuje. Například pro cílenou metabolomiku se používají obvykle kvadrupólové hmotnostní spektrometry pracující s jednotkovým rozlišením, které na základě monitorování vybraných reakcí (angl. „selected reaction monitoring“, SRM) v analyzovaném vzorku detekují předem definované analyty. Proto se pomocí různých extrakčních postupů vzorek připraví tak, aby cílová skupina analytů mohla být změřena co nejpřesněji (potlačení analytických artefaktů a vlivu látek přítomných ve vysoké koncentraci) [7,20,21].

V případě necílené metabolomiky se pro detekci používá hmotnostní spektrometrie s vysokým rozlišením [6,22], s cílem zachytit ve vzorku pokud možno komplexní profil přítomných analytů (včetně neznámých) s vysokou přesností a správností určení poměru hmotnosti a náboje ( $m/z$ ) [7,22]. Vzorky se připravují tak, aby se jejich metabolom změnil co nejméně [23], proto je třeba používat postupy umožňující rychlé zastavení metabolických změn (např. prudké zchlazení při současné deproteinaci). Vzhledem ke složení metabolomu je při měření nezbytný velký dynamický rozsah analytických systémů. Jelikož je ionizace elektrosprejem v iontovém zdroji hmotnostního spektrometru kompetitivním procesem, je silně závislá na komplexnosti vzorku. Během ionizace analytu v přítomnosti interferentů (typicky lipidy či soli) dochází ke kompetici o náboj z elektrospreje, k poklesu ionizační účinnosti a následně pak potlačení signálu v hmotnostním analyzátoru. Tento jev se nazývá iontová suprese [24]. Díky tomuto bývá instrumentálně náročné zachytit stopová množství látek v přítomnosti jiných analytů o vysoké koncentraci (a obecně problém s kvantifikací).

Druhým velkým úskalím necíleného přístupu je analýza velkého množství dat, které tento přístup generuje. Jejich zpracování vyžaduje mnohem pokročilejší chemometrické nástroje, vysoký početní výkon a čas [22,25]. Jednotlivé analyty se identifikují pomocí spektrálních knihoven nebo dalšími analytickými postupy [25]. Analytický signál však neposkytují jen analyty, ale i řada artefaktů,

keré se do datových souborů dostávají při přípravě a analýze vzorku (např. ftaláty, polyethylenglykol, thiodiglykol). Tyto kontaminanty lze odhalit prostřednictvím analýzy slepého vzorku - blanku (vzorek, který je zpracováván stejně jako experimentální vzorky; biologický materiál je v něm nahrazen odpovídajícím množstvím vody) a statistických metod (artefakty jsou v podobné koncentraci přítomny ve všech vzorcích, a tak nejsou hodnoceny jako statisticky významné) [8]. Tento postup nicméně funguje pouze v případě, že všechny vzorky včetně slepých byly odebrány a připraveny zcela identicky, za použití stejné šarže chemických reagensů a spotřebního materiálu.

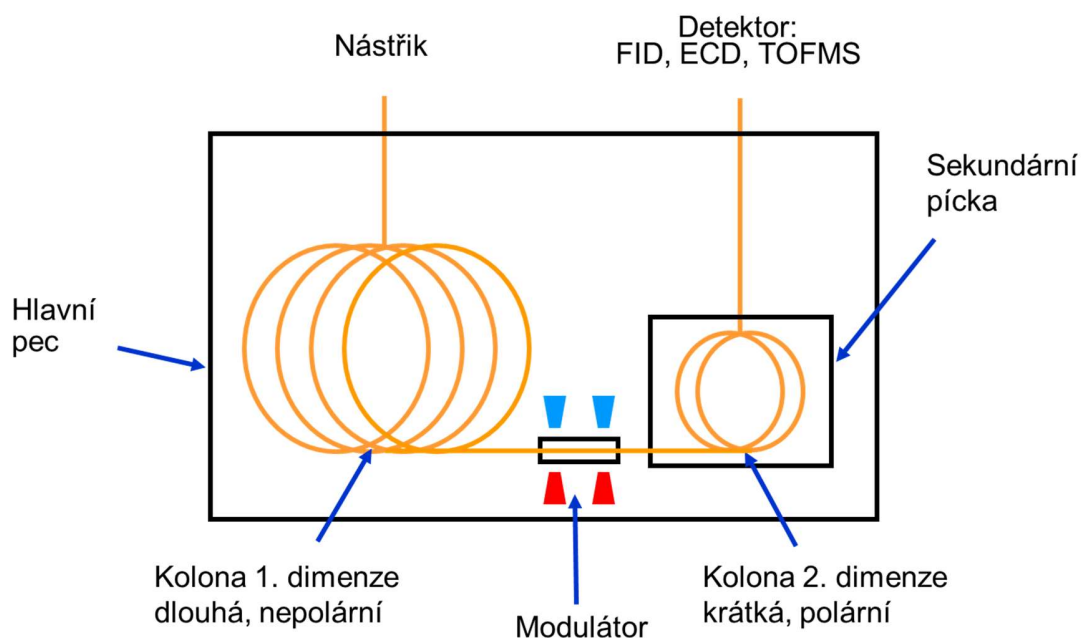
## 4 Instrumentální analytické techniky v metabolomice

Velká chemická různorodost cílových látek klade vysoké nároky na analytické nástroje používané v metabolomice [10]. Důležitá je dostatečná citlivost ( $< \mu\text{mol.l}^{-1}$ ), velký dynamický rozsah měření (rozdíly v koncentracích mezi jednotlivými analyty jsou až 12 řádů [26]), vysoká frakcionační kapacita (některé ze sledovaných metabolitů jsou si strukturně velmi podobné) a dostatečná stabilita systému (často se měří série až tisíců vzorků [27]).

V současné době neexistuje analytický systém, který by byl sám schopen měřit celý metabolom [17]. Proto se metabolomické studie zaměřují buď na určitou část metabolomu (např. organické kyseliny) nebo se použije kombinace několika technik. Pro akvizici dat se většinou používá nukleární magnetická rezonance (NMR) nebo hmotnostní spektrometrie (MS). Zatímco NMR vyniká vysokou reprodukovatelností výsledků a minimální přípravou vzorků [5], MS citlivostí [10]. Frakcionační kapacita MS je však velmi nízká, proto bývá spojována s některou ze separačních technik: nejčastěji s kapalinovou chromatografií (LC), méně pak s plynovou chromatografií (GC) a zřídka i s kapilární elektroforézou [10]. Nevýhodou GC je zdlouhavá příprava vzorku a vyšší detekční limit (řádově v  $\mu\text{mol.l}^{-1}$ ) [20], u LC se jedná o nízkou reprodukovatelnost retenčních časů, nižší životnost kolon a separační účinnost.

Všechny výše uvedené analytické techniky lze označit jako vícerozměrné separační techniky, jelikož spojují dva separační systémy s různým principem separace (např. GC a MS). Při analýze komplexních vzorků se ve spojení s MS s výhodou používají i dvojrozměrné chromatografické systémy, ve kterých separace v každé z dimenzí probíhá na základě jiného mechanismu, čímž se mnohonásobně navýší frakcionační kapacita [28,29]. Příkladem takového systému je dvojdimenzionální plynový chromatograf (GC×GC, Obr. 3), který obsahuje dvě sériově uspořádané kolony o různé selektivitě spojené prostřednictvím modulátoru [30]. Modulátor nastříkuje efluent z první dimenze do druhé tak, aby byla zachována separace v dosažená v koloně první dimenze [31]. Modulátor proto musí pracovat

velmi rychle (obvykle s frekvencí 1-4 Hz), kolona druhé dimenze musí být kratší než první (30-60m vs. 1-2m) [32]. Také detektor, nejčastěji TOF MS (hmotnostní spektrometr obsahující hmotnostní analyzátor doby letu – z anglického „time of flight“), zřídka i plamenný ionizační detektor (FID) či detektor elektronového záchytu (ECD), musí mít vysokou rychlost sběru dat (až 500 Hz), aby zajistil dostatečnou hustotu dat i pro píky s velmi vysokou separační účinností, typické pro GC×GC (šířka píku v polovině výšky 0,05 s). To, zda takový systém představuje benefit pro analýzu konkrétního typu vzorku, se ověřuje mnoha způsoby. Mezi nejznámější patří stanovení míry nárůstu píkovej kapacity [33] a určení ortogonality [31]. V této práci je využito ortogonálního přístupu.



Obr. 3: Obvyklé uspořádání dvojdimenzionálního plynového chromatografu

#### 4.1 Výpočet ortogonality dvojrozměrných chromatografických systémů

Ortogonalita ( $O$ ) je parametr, který hodnotí, do jaké míry jsou dimenze daného vícedimenzionálního separačního systému odlišné – zda každá z dimenzí k separaci využívá jinou vlastnost daného vzorku [31]. Obecně lze říci, že ortogonalita vypovídá o tom, zda připojení další dimenze přináší nějakou novou informaci o vzorku. Ortogonalita může nabývat hodnot od 0 do 100 %. Čím je ortogonalita vyšší, tím je vícerozměrný analytický systém pro danou analýzu přínosnější.

Existuje několik desítek způsobů určení ortogonality, které lze podle základního principu rozdělit do tří skupin. První skupina stanovuje, jak moc jsou si informace přinášeny první a druhou

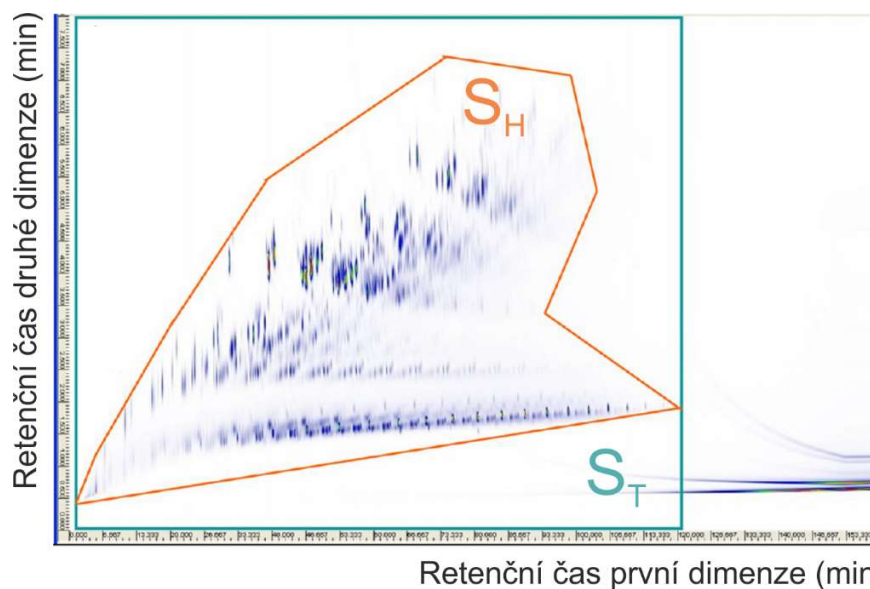
dimenzí podobné. Patří sem korelační analýza a přístupy založené na informační teorii (např. podmíněná entropie). Korelační analýza (např. Pearsonova korelace, PC) hodnotí matematickou závislost mezi retenčními časy první a druhé dimenze. Čím jsou si podobnější, tím je hodnota korelačního koeficientu vyšší. Ortogonalita se pak spočítá inverzí absolutní hodnoty korelačního koeficientu ( $r$ ) podle rovnice (Rov. 1, [34]). Vysoce korelované dimenze proto poukazují na nízkou efektivitu separace (ortogonalita systému se blíží nule) a naopak.

$$O = (1 - |r|) \cdot 100 \% \quad (\text{Rov. 1}),$$

kde  $r$  je korelační koeficient.

Podmíněná entropie (CE) počítá na základě retenčních dat první dimenze teoretická retenční data dimenze druhé. V podobě veličiny  $O$  pak vyjadřuje míru shody těchto teoreticky vypočítaných dat s daty reálně změřenými druhou dimenzí. Čím více se shodují, tím je ortogonalita systému a tím i separační účinnost vyšší [35].

Druhá skupina ortogonálních přístupů popisuje kvalitu pokrytí separačního prostoru. Konvexní obálka (CH, Obr. 4) obkreslí pomocí Delaunayova algoritmu plochu obsazenou píky ( $S_H$ ) a porovná jí s celkovou využitelnou plochou separačního prostoru ohraničeného dvěma dimenzemi ( $S_T$ , Rov. 2, [36]).

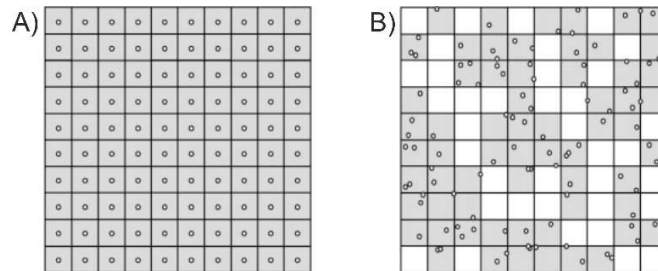


Obr. 4: Výpočet využití separačního prostoru pomocí metody konvexní obálky:  $S_H$  - plocha obsazená píky,  $S_T$  - celková využitelná plocha separačního prostoru ohraničeného dvěma dimenzemi. GC×GC chromatogram separace solventní nafty, vypočítaná ortogonalita systému 54,5 %. Převzato z [37].

$$O = \frac{S_H}{S_T} \cdot 100 \% \quad (\text{Rov. 2}),$$

kde  $S_H$  - plocha obsazená píky vymezená konvexní obálkou,  $S_T$  - celková využitelná plocha separačního prostoru ohraničeného dvěma dimenzemi.

Přístup založený na geometrickém pokrytí separačního prostoru (SC, Obr. 5) rozděluje separační prostor do tolika stejně velkých čtverců („binů“) jaká je nejbližší celočíselná hodnota odmocniny z celkového počtu píků (např. pro 100 píků je to 10, pro 148 píků 12). Pokud je každý píků umístěn ve svém čtverci, je systém dokonale ortogonální (Obr. 5A). Tento stav je však dosažitelný pouze teoreticky. V praxi je nutno započítat i Poissonovu konstantu pravděpodobnosti toho, že se píků náhodně trefí právě do svého binu (Obr. 5B), která pro 100 píků činí 63 % [38]. Ortogonalita se pak spočítá podle rovnice 3 (Rov. 3).



Obr. 5: Výpočet ortogonalit dle přístupu geometrického pokrytí separačního prostoru: A) Teoretické maximální pokrytí, B) Praktické maximální pokrytí po zahrnutí Poissonova rozdělení pravděpodobnosti náhodných jevů. Převzato z [38].

$$O = \frac{\sum Bin}{0,63 \cdot \sum P} \quad (\text{Rov. 3}),$$

kde  $\sum Bin$  je počet binů obsazených alespoň jedním píků,  $\sum P$  je celkový počet binů v separačním prostoru.

Hodnota 0,63 však není konstantou, její velikost závisí na počtu píků v analýze. Watson a kol. [39] prostřednictvím 2000 simulací prokázali, že pro 100 píků se její hodnota pohybuje v rozmezí od 0,51 do 0,74 (tj. 0,63 v průměru), zatímco pro 2500 píků od 0,61 do 0,66 (tj. 0,64 v průměru). Pro nižší počet píků v analýze nabývá konstanta nižších hodnot (např. pro 5 a 10 píků průměrně 0,47 resp. 0,57). Pokud je v analýze více píků než binů (jako např. ve výše uvedeném případě, kdy je v analýze 148 píků a separační prostor je rozdělen na 12×12 binů), vzniká další nepřesnost (4 píky, které jsou navíc, mohou

vyplnit 4 biny a uměle navýšit ortogonalitu). Rovnici 3 je proto nutné používat v korigované formě (Rov. 4).

$$O = \frac{\sum Bin}{(1-e^{-\alpha}) \cdot \sum P} \quad (\text{Rov. 4}),$$

kde  $\sum Bin$  je počet binů obsazených alespoň jedním píkem,  $\sum P$  je celkový počet binů v separačním prostoru a  $\alpha$  je poměr píků k celkovému počtu binů.

Metoda nejbližšího souseda (NN) měří pro každý píku, jak je od něj v separačním prostoru vzdálený jeho nejbližší soused [40]. Kvalitu separace pak popisuje průměrná vzdálenost všech píků od jejich nejbližších sousedů. Čím vyšší je absolutní hodnota průměru, tím lepší je separace. Způsob výpočtu průměru nejbližších vzdáleností zohledňuje zvolený aspekt separace. Pokud se použije harmonický průměr (HN, Rov. 5), je větší důraz kladen na krátké vzdálenosti. Prostřednictvím HN se proto hodnotí rovnoměrnost pokrytí separačního prostoru (nepřítomnost shluků píků). Tento aspekt separace je označován jako lokální ortogonalita.

$$HN (di_1, di_2, \dots, di_n) = \frac{n}{\frac{1}{di_1} + \frac{1}{di_2} + \dots + \frac{1}{di_n}} \quad (\text{Rov. 5}),$$

kde  $di$  je vzdálenost daného píku od jeho nejbližšího souseda.

Pokud se použije aritmetický průměr (AN, Rov. 6), jsou naopak zohledňovány velké vzdálenosti. Prostřednictvím AN se tak hodnotí rozsah využitého separačního prostoru. Tento aspekt separace je označován jako globální ortogonalita.

$$AN (di_1, di_2, \dots, di_n) = \frac{di_1 + di_2 + \dots + di_n}{n} \quad (\text{Rov. 6}),$$

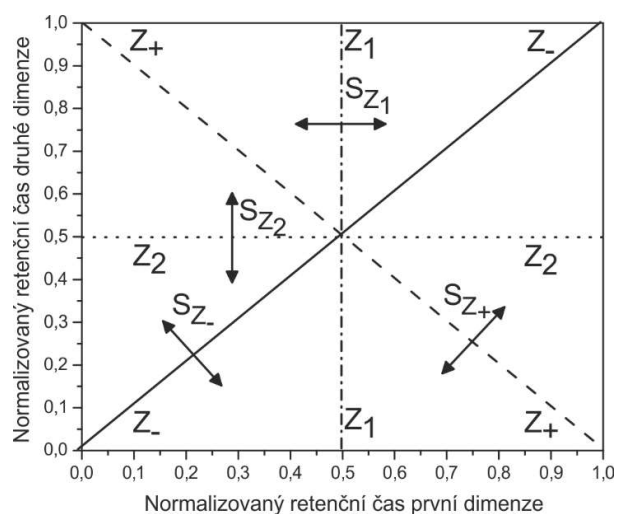
kde  $di$  je vzdálenost daného píku od jeho nejbližšího souseda.

Při použití přístupu založeného na tzv. hvězdičkových rovnicích (AE) je ortogonalita systému popsána jako distribuce píků podél čtyř os  $Z_+$ ,  $Z_-$ ,  $Z_1$ ,  $Z_2$  (Obr. 5). Pro každou z os se počítá průměrná vzdálenost všech píků od dané osy [41]. Ortogonalita systému je pak rovna druhé odmocnině z násobku hodnot poskytnutých jednotlivými rovnicemi (Rov. 7).

$$O = \sqrt{Z_+ \cdot Z_- \cdot Z_1 \cdot Z_2} \quad (\text{Rov. 7}),$$

kde  $Z_+$ ,  $Z_-$ ,  $Z_1$ ,  $Z_2$  jsou distribuce píků podél stejnojmenných os, vyjádřené jako průměrná směrodatná odchylka všech bodů v separaci od hodnoty představované danou osou (např. pro osu  $Z_1$  a  $Z_2$  je to hodnota 0,5, viz Obr. 6).





Obr. 6: Výpočet ortogonalit systému pomocí hvězdičkových rovnic. Převzato z [41].

Třetí skupina ortogonálních přístupů zahrnuje metody, které v sobě kombinují přístupy obou předchozích skupin (pokrytí separačního prostoru a podobnost informací přinášejících první a druhou dimenzí). Jedná se např. o modelování (MA), ve kterém se hodnota získaná pomocí algoritmu pro geometrické pokrytí separačního prostoru vynásobí inverzní hodnotou koeficientu determinace ( $r^2$ , Rov. 8, [42]).

$$O = \frac{\sum Bin}{(1 - e^{-\alpha}) \cdot \sum P} \cdot (1 - r^2) \quad (\text{Rov. 8}),$$

Kde  $\sum Bin$  je počet binů obsazených alespoň jedním píkem,  $\sum P$  je celkový počet binů v separačním prostoru a  $\alpha$  je poměr piků k celkovému počtu binů,  $r^2$  je koeficient determinace.

## 4.2 Škálování separačního prostoru

V jednodimenzionální chromatografii je záznamem chromatografického procesu lineární chromatogram, kdy na ose  $x$  je retenční čas, na ose  $y$  je intenzita měřeného signálu. Ve dvoudimenzionální chromatografii mluvíme o separačním prostoru, který je vymezen retenčním časem první dimenze na ose  $x$ , retenčním časem druhé dimenze na ose  $y$  a intenzitou na ose  $z$ . Rozsah první a druhé dimenze se často výrazně odlišuje, zejména v GCxGC (kde je u první dimenze je např. 0-3000 s, zatímco u druhé 0-4 s). Aby byl kladen stejný důraz na separaci první a druhé dimenze, data se před výpočtem ortogonalit škálují. Vzniká tak normalizovaný retenční prostor (proto se tento krok nesprávně označuje i jako normalizace), který má v obou dimenzích rozsah od 0 do 1 (viz Obr. 6).

Zároveň se z chromatogramu odstraní nevyužitelný (nebo nevyužitý – zde se názory různí – viz diskuse) prostor (viz Obr. 4, kde je využitelný prostor ohraničen zeleně).

Škálování se provádí pomocí dvou různých algoritmů. První z nich lze označit jako „void-max“ (z anglického „void time“ – mrtvý čas a „maximal retention time“ – nejvyšší retenční čas), protože dvojdimenzionální separační prostor je v každé z obou dimenzí ohraničen mrtvým časem dané dimenze ( $t_0$ ) a maximálním retenčním časem analytu, kterého bylo v dané dimenzi při dané separaci dosaženo ( $t_{max}$ ) [43]. Toto škálování je ze všech nejstarší, často (zvláště ve starších pracích) bývá označováno i jako normalizace podle průměrné retence [44]. Retenční časy se při jeho aplikaci přepočítají podle rovnice 9:

$$X(A) = \frac{t_r(A) - t_0}{t_r(max) - t_0} \quad (\text{Rov. 9}),$$

kde  $X(A)$  je normalizovaný retenční čas látky A,  $t_r(A)$  je retenční čas látky a v dané dimenzi,  $t_0$  je mrtvý čas dané dimenze,  $t_r(max)$  je maximální retenční čas, kterého bylo v dané dimenzi dosaženo.

Druhý algoritmus lze označit jako „min-max“ (z anglického „minimal retention time“ – nejnižší retenční čas a „maximal retention time“ – nejvyšší retenční čas). Při jeho aplikaci se retenční časy první a druhé dimenze škálují podle retenčního času nejméně a nejvíce zadržované látky (Rov. 10).

$$X(A) = \frac{t_r(A) - t_r(min)}{t_r(max) - t_r(min)} \quad (\text{Rov. 10}),$$

kde  $X(A)$  je normalizovaný retenční čas látky A,  $t_r(A)$  je retenční čas látky A v dané dimenzi,  $t_r(min)$  je retenční čas nejméně zadržovaného analytu,  $t_r(max)$  je retenční čas nejvíce zadržovaného analytu.

### 4.3 Dimenzionalita vícerozměrných separací

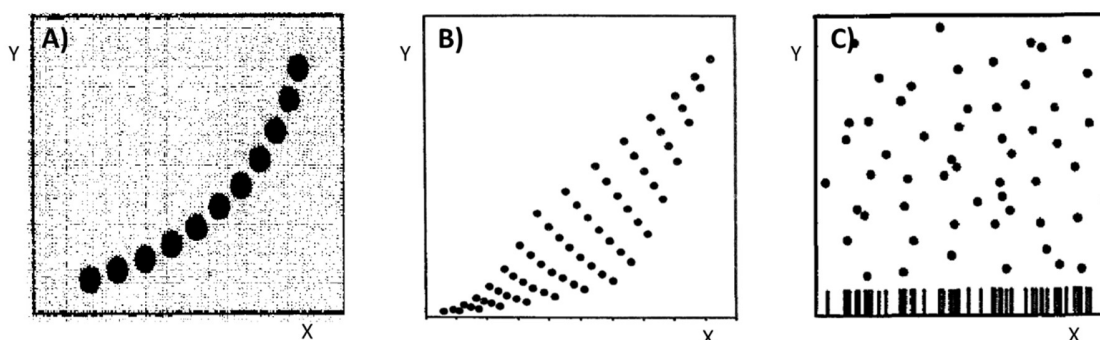
U vícerozměrných separací se setkáváme se dvěma druhy dimenzionality: dimenzionalitou vzorku [45] a dimenzionalitou separačního (chromatografického) systému [30]. Oba druhy dimenzionality spolu úzce souvisí, neboť nejvíce analytických informací o vzorku získáme tehdy, když se dimenzionalita vzorku a systému shoduje [45].

Dimenzionalita analytického systému je dána počtem spřažených separačních dimenzí. Nárůst frakcionační kapacity systému jde však na vrub jednoduchosti datového výstupu. Proto běžné analytické systémy mají dimenzionalitu zřídka vyšší než 2.

Dimenzionalita vzorku je souhrnem všech vlastností analytů vzorku, které mohou přispět k jeho chromatografickému chování [45]. Jednodimenzionální vzorky jsou tvořeny výhradně látkami stejné

chemické třídy, které jsou se stacionární fází schopny interagovat pouze pomocí jediného mechanismu. Příkladem jednodimenzionálního vzorku je směs alkanů. K jejich separaci dochází výhradně na principu Londonových disperzních sil. Použití vícedimenzionálního systému nemá pro jejich analýzu žádný přínos, protože přidání další separační dimenze nemůže o vzorku přinést žádnou novou informaci (analyty vzorku s ní buď interagují stejně jako s první dimenzí nebo s ní neinteragují vůbec). Analyty takovýchto vzorků tvoří na 2D chromatogramu úzký pás (buď ve formě úhlopříčky (Obr. 7A) nebo rovnoběžný s jednou z os).

Dvojdimenzionální vzorky (např. směs alkoholů) jsou schopny se rozdělit na základě dvou různých separačních mechanismů (Londonovy disperzní síly v první dimenzi a vodíkové vazby v dimenzi druhé). Takovéto vzorky, které mají stejný počet dimenzí jako použitý chromatografický systém, se označují jako nízkodimenzionální [45]. Poskytují strukturované chromatogramy (Obr. 7B), které lze s výhodou použít k identifikaci neznámých látek [46], neboť analyty homologních sérií v nich tvoří skupiny nebo linie [45].



Obr. 7: Typický vzhled 2D chromatogramů A) jednodimenzionálních, B) dvojdimenzionálních, C) vysokodimenzionálních vzorků při separaci dvojdimenzionálním separačním systémem (převzato z [30,45]). X, Y – normalizovaný retenční čas první a druhé dimenze.

Většina reálných vzorků však patří mezi vzorky vysokodimenzionální, jejich dimenzionalita je až o tři řády vyšší než dimenzionalita separačního systému [45]. Typickým příkladem je metabolom moči, který obsahuje tisíce látek [13] různých chemických tříd (m.j. organické kyseliny, aminokyseliny, puriny, pyrimidiny, acylkarnitiny, acylglyciny, polyoly). Funkční skupiny těchto látek různou měrou interagují se stacionární fází první a druhé dimenze, proto chromatogramy vysokodimenzionálních vzorků postrádají vnitřní strukturu, celý 2D separační prostor je pokryt píky analytů (Obr. 7C). Z Giddingsova [45] odhadu 10 000 různých interakcí analytů vysokodimenzionálních vzorků se stacionární fází je zřejmé, že interpretace jejich chromatogramů je poměrně obtížná. Naštěstí ne všechny funkční skupiny analytů

interagují se stacionární fází stejnou měrou. Ve vzorku je celá řada takových, které se stacionární fází interagují jen velmi slabě. Na chromatogramu se tyto interakce téměř neprojeví. Vzorek se pak chromatograficky chová tak, jako by jeho dimenzionalita byla mnohem nižší. Tohoto jevu, jemuž se říká „zjevná dimenzionalita“ (z angl. „apparent dimenzionalita“), lze využít k analýze komplexních vzorků. Při ní se vzorek upraví tak, aby byla jeho dimenzionalita co nejnižší, ale stále taková, aby došlo k dosažení cíle analýzy. Tomuto druhu dimenzionality se říká „požadovaná dimenzionalita“ (z angl. „required dimenzionalita“). Rutinně se už mnoho let využívá v aplikacích, kde cílem analýzy je pouze určitá (v biomedicíně např. diagnosticky významná) skupina látek. Příkladem je analýza organických kyselin v moči, při které se používá kyselá extrakce do ethylacetátu. Díky ní jsou pak analyzovány pouze látky, které při pH 1-2 přejdou do organické fáze [20]. Všechny ostatní látky přítomné v moči nejsou při diagnostice organických acidurií podstatné a mohou být opomenuty.

## 5 Sběr a příprava vzorků

Správný počet vzorků se pro konkrétní metabolický experiment zvolí tak, aby byla dostatečná síla studie [47]. Skupina vzorků pacientů musí být pokud možno homogenní (např. stejný typ nádoru nebo stejná metabolická porucha, obdobná nutriční, nepřítomnost infekce, podobný věk a pohlaví). Důležitý je i výběr vzorků pro kontrolní skupinu. Kontrolní vzorky by měly být sbírány ve stejnou dobu, na stejném místě, do stejných odběrových nádob a skladovány stejně, jako vzorky pacientů. Výběr těchto vzorků také musí respektovat povahu studie, aby se např. nestalo, že kontrolní skupinou pro nalezení biomarkerů pro diagnostiku karcinomu prostaty budou jedinci s opačným pohlavím a jiným věkem [48]. Kontrolní subjekty by obecně měly být stejného věku, pohlaví a nutriční jako studované subjekty [49].

V metabolomice se velký důraz klade na preanalytickou fázi, neboť sledované metabolity nebývají vždy stabilní [50]. Kritickým obdobím bývá časový interval od samotného odběru do zamrazení vzorku [51] a dále pak i vlastní příprava vzorku k analýze [49]. Obecně mají být tyto intervaly byly co nejkratší a pro všechny vzorky stejné. Důležité je také správné skladování vzorku. Vzorky se obvykle skladují při -80 °C [27]. Ale ani tyto podmínky nejsou vždy dostatečné pro zabránění změn metabolomu především při dlouhodobém skladování vzorků. Vlastní příprava vzorku se odvíjí od povahy studovaných metabolitů, analyzovaného materiálu, zvoleného metabolického přístupu, analytické techniky a metody analýzy [11].

Pro analýzu vzorků plasmy a séra pomocí LC-MS se příprava obvykle sestává pouze z odstranění proteinů precipitací (tzv. deproteinace). Volba precipitačního činidla je rozhodující pro spektrum

analytů přítomných ve vzorku [52–57]. Například při použití směsi methanolu s ethanolem v poměru 1:1 společně s proteinasou K místo nejčastěji používaného 80% methanolu se významně zvýší obsah hydrofobních metabolitů ve vzorku [53]. Při zpracování buněčných kultur se buňky omyjí od zbytku kultivačního media, zastaví se jejich metabolismus (tzv. quenching), nakonec se rozruší jejich membrány a uvolní buněčný obsah [58,59]. Moč představuje ve srovnání s plasmou a buňkami méně komplexní biologický materiál, navíc s velmi malým obsahem proteinů [60]. Proto se při přípravě pro kapalinovou chromatografii často pouze ředí na stejnou koncentraci kreatininu ve všech vzorcích (korekce na různou hydrataci organismu) [15] a následně filtruje.

Přípravu vzorků moči pro plynovou chromatografii komplikují dva jevy: kapalně skupenství vzorků (separace probíhá v plynné fázi) a přítomnost močoviny (coby nejvíce zastoupené složky matrice). Převod analytů do plynného skupenství usnadňuje derivatizace. Derivatizačních postupů je známa celá řada. Nejpoužívanější je trimethylsilylace, která zvyšuje těkavost sloučenin tím, že nahrazuje kyselé vodíky, schopné tvořit vodíkové vazby, trimethylsilylovou skupinou [20,61]. Při použití této metody však není možné od sebe rozlišit ketony a enoly. Proto se před trimethylsilylací provádí ještě oximace, která fixuje ketony tím, že kyslík nahradí dusíkem, na nějž je přes kyslík navázán substituent (nejčastěji methyl- či ethyl- [62]).

Močovina je ze vzorků odstraňována zejména dvěma různými způsoby: buď pomocí kyselé extrakce do organického rozpouštědla, kam jako kladně nabitá látka přechází jen v omezené míře [20] (viz kapitola 4.3), a nebo za použití ureasy [63–65]. Ureasa je enzym, který z endogenních metabolitů štěpí pouze močovinu za vzniku amoniaku a vody. Hladinu ostatních metabolitů ureasa ovlivňuje pouze nepřímo. Nejprve při vlastní enzymatické reakci, kde působí změnu pH a tím ovlivňuje rozpustnost analytů, pak při jejím odstraňování z reakční směsi (nejčastěji deproteinací ledovým methanolem). Methanol působí částečně jako extrakční činidlo, a tak mění profil metabolitů [66]. Navíc po deproteinaci zůstává ve vzorku řada artefaktů, které komplikují vlastní GC-MS analýzu [63].

Sérum [67] i plasma [68] se obvykle deproteinují methanolem a stejně jako buněčný obsah [59,69] se před GC-MS analýzou odfoúkají pod proudem dusíku, případně lyofilizují. Následuje dvoustupňová derivatizace (oximace, trimethylsilylace).

Vzorky se měří v rámci jedné sekvence, a to v náhodném pořadí (angl. „randomized order“) [7]. S každou sérií vzorků jsou vždy analyzovány blanky a také kontrolní vzorky (označované jako „QC“, z anglického „quality control“), které sledují stabilitu použitého analytického systému [27,70]. Každý QC je připraven smísením alikvotního podílu každého vzorků daného experimentu. Měřený signál analytů se často mění v čase. Vzorky měřené na začátku experimentu často poskytují vyšší odezvu než vzorky

měřené na jeho konci. Je-li v průběhu experimentu v pravidelných intervalech měřen stejný vzorek (QC), lze odezvy analytů ve vzorcích na základě jejich odezvy v QC matematicky korigovat polynomem (pomocí techniky LOESS – z anglického „locally estimated scatterplot smoothing“) [27].

## 6 Vyhodnocování metabolomických dat

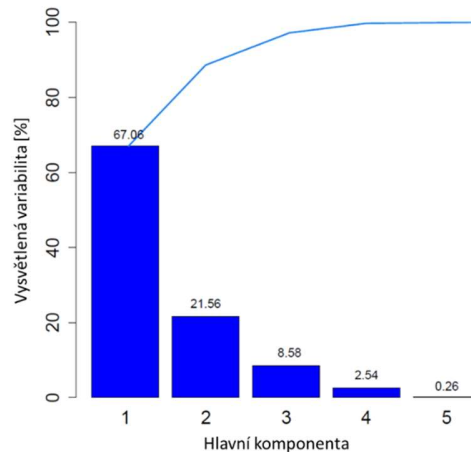
Datové soubory vyhodnocované v metabolomických studiích jsou komplexní, obsahující stovky proměnných (abundance jednotlivých metabolitů) od desítek až stovek pozorování (vyšetřovaných subjektů - vzorků) [71]. Nelze je tedy jednoduše vyhodnotit. Navíc surová data poskytovaná analytickými systémy jsou zatížena řadou jevů, jako je nestabilita analytického systému během měření, přítomnost artefaktů, vznik více forem stejného metabolitu (např. u GC-MS obvykle izomery, derivatizační produkty, izotopy, fragmenty). Často také (jako typická biologická data) nemají normální rozdělení. Proto vlastnímu vyhodnocení dat předchází jejich předběžná analýza a úprava [58]. Patří do ní např. nastavení základní linie hladiny šumu, správné zarovnání píků analytů (angl. „peak alignment“), nahrazení chybějících hodnot, nastavení parametrů integrace píků [72,73]. Naměřená data (plochy píků) se pak pomocí LOESS (viz kapitola 3) upraví tak, jako by analytický systém poskytoval stabilní odezvu. Analyty, které v QC vzorcích i po tomto kroku vykazují variabilitu vyšší než 30 %, jsou vyloučeny z další analýzy [74]. LOESS navíc data i škáluje (převede abundanci proměnných na stejné měřítko).

Dále se data upravují transformací: nejčastěji pomocí *clr* (z anglického „centered log ratio“, [21,75–77]), aplikací PQN (z anglického „probabilistic quotient normalization“, [72,78]) či tradičním logaritmizací [79], aby získala normální rozdělení [80]. Další úpravou dat, která se před použitím vícerozměrných analýz provádí, je centrování dat. Při centrování je od původní hodnoty odečítán průměr proměnné, čímž zajistí, že střední hodnotou, kolem které osciluje sledovaná proměnná, je nula [58]. Následuje průzkumová analýza dat, která zahrnuje testování datového souboru na normalitu, korelace, homogenitu a určení odlehlých hodnot [81]. Tento krok je nezbytný pro správný výběr statistických testů. Pokud mají data normální rozdělení, lze použít parametrické testy (např. t-test), v opačném případě je nutné použít neparametrické testování (např. Mann-Whitneyův U-test).

Vlastní analýza dat metabolomického experimentu je pak kromě jednorozměrných statistických metod provedena vícerozměrnými statistickými metodami, a to ve čtyřech krocích:

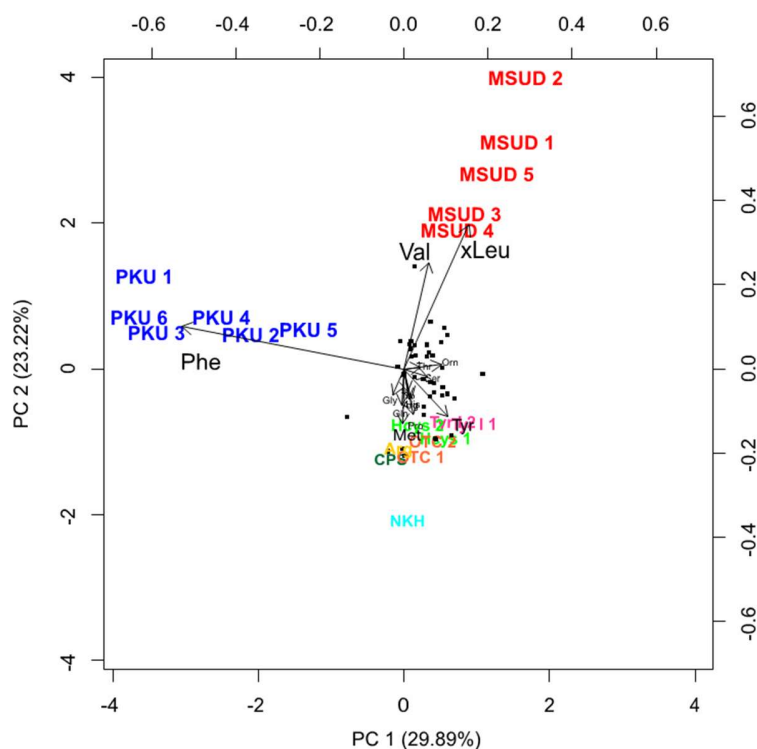
- 1) Nalezení takových kombinací původních proměnných, které co nejlépe vystihují proměnlivost původních dat. Tyto nové proměnné se označují jako hlavní komponenty, jsou na sebe kolmé a vznikají lineární kombinací původních dat. K interpretaci se většinou

používají první dvě, které vystihují nejvíce variability v datech (Obr.8). Navíc umožňují zobrazit data v rovině [81].



Obr. 8: Indexový graf úpatí vlastních čísel (angl. „scree plot“) zobrazující hlavní komponenty a procenta vysvětlené variability. První dvě komponenty vysvětlují 88,62 % variability. Vlastní data (stanovení nejhodnějšího algoritmu pro výpočet ortogonalitu 2D analýzy nízkodimenzionálních vzorků).

2) Hledání vztahů mezi proměnnými (třídění do shluků podle podobnosti). Pro odhalení trendu v datech se použije neřízená (algoritmická) metoda (angl. „unsupervised“), např. analýza hlavních komponent (PCA), která nemá informace o příslušnosti vzorků k jednotlivým skupinám (např. pacientů a zdravých kontrol). Jedná se vlastně o průzkumovou analýzu vícerozměrných dat [82]. Grafickým výstupem PCA je obvykle rozptylový diagram komponentních skór (angl. „score plot“), graf komponentních zátěží (angl. „loading plot“), případně dvojný graf (angl. „biplot“, Obr. 9), který kombinuje oba předchozí grafy [58]. Na rozptylovém diagramu komponentních skór je každý subjekt (pacient, kontrola) vyjádřen jako bod (charakterizující všechny proměnné = analyty daného subjektu) promítnutý do roviny zobrazených hlavních komponent. Subjekty se podobnými profily analytů tvoří na rozptylovém diagramu komponentních skór shluky. Graf komponentních zátěží pomocí vektorů („šipek“), které mají počátek v průsečíku obou hlavních komponent, ukazuje, do jaké míry daný analyt ovlivňuje hlavní komponenty. Čím delší je daný vektor, tím větší vliv daný analyt na hlavní komponenty má. Úhly, které mezi sebou vektory svírají, vyjadřují vztahy mezi analyty: jsou-li vektory víceméně souběžné (úhel je malý), jedná se o pozitivně korelované analyty, svírají-li pravý úhel, jde o nekorelované analyty, svírají-li úhel okolo 180 °, jedná se o negativně korelované analyty.

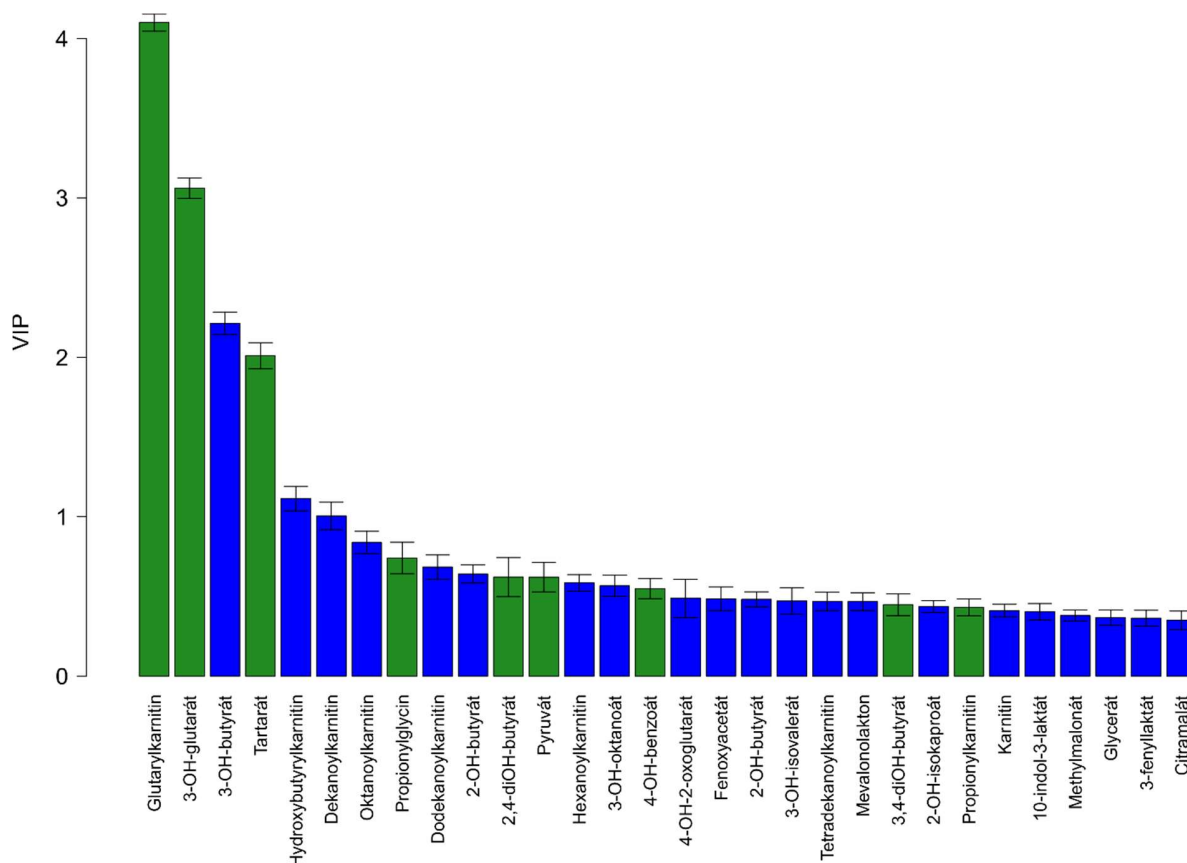


Obr. 9: Grafický výstup PCA: dvojný graf jako kombinace rozptylového diagramu komponentních skóre (osa vlevo a dole) a grafu komponentních zátěží (osa vpravo a nahoře). Číselné hodnoty vedle popisu os udávají procenta vysvětlené variability danou komponentou. Jedná se o cílenou metabolickou analýzu aminokyselin ve vzorcích plasmu pacientů s různými metabolickými poruchami. Převzato z [21]. Rozptylový diagram komponentních skór ukazuje, že vzorky pacientů se stejným onemocněním (např. fenylketonurií, PKU) se shlukují a že vzorky pacientů s PKU a leucinózou (MSUD) se od ostatních nejvíce liší. Graf komponentních zátěží ukazuje, že vzorky s PKU nejvíce odlišuje hladina fenylalaninu (Phe), zatímco v případě MSUD je to součet leucinu a izoleucinu (xLeu) a valin (Val).

3) Hledání rozdílů mezi skupinami. Rozdíly mezi skupinami jsou stanoveny pomocí diskriminační analýzy na základě metody částečných nejmenších čtverců (PLS-DA, z angl. „partial least squares discriminant analysis“ [76,83]). PLS-DA promítne pomocí metody nejmenších čtverců data do nových (tzv. latentních) proměnných a tím i do nového (lineárního) prostoru [84]. Diskriminační analýza potom ověří, zda se od sebe předem definované skupiny (většinou jen 2, např. pacienti a kontroly) skutečně odlišují a identifikuje proměnné (analyty), které nejlépe popíší rozdíly mezi nimi. Také najde celkový trend pro danou skupinu bez ohledu na interindividuální rozdíly mezi subjekty v rámci skupiny [83]. Jedná se o metodu řízenou (heuristickou, angl. „supervised“), na jejímž začátku je nutno zadat příslušnost každého vzorku k té které skupině. Grafické výstupy PLS-DA jsou obdobné jako u PCA:



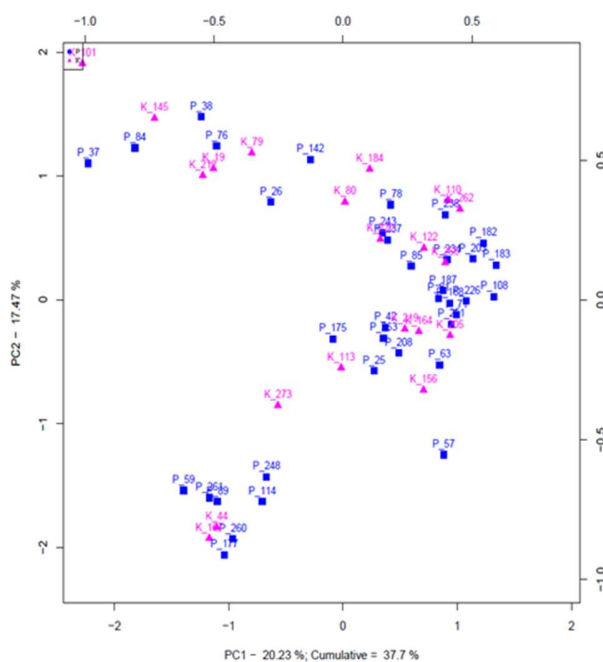
rozptylový diagram komponentních skóre, graf komponentních zátěží a dvojný graf. Navíc se zde vyskytuje graf významnosti proměnných v projekci použité pro PLS-DA model (VIP, z angl. „variable importance in the projection“) ukazující VIP skóre jednotlivých proměnných (analytů, Obr. 10) a vypovídá o tom, jak moc daná proměnná přispívá k tvorbě modelu [85,86]. Všechny analyty s  $VIP > 1$  jsou ze statistického hlediska signifikantní [85,87]. Toto kritérium však nelze zobecnit, protože je závislé na povaze konkrétních dat [88]. Z biologického hlediska jsou za signifikantně zvýšené či snížené považovány proměnné s  $VIP > 2$  [89].



Obr. 10: VIP graf zobrazující VIP skóre proměnných PLS-DA modelu. LC-MS/MS analýza organických kyselin v séru pacientů s glutarovou acidemií I. typu a zdravých kontrol (vlastní nepublikovaná data). Analyty zvýšené u pacientů s GAI jsou vyobrazeny zeleně, analyty zvýšené u zdravých kontrol modře. Analyty s  $VIP > 2$  jsou z biologického hlediska signifikantní: glutarylkarmitin, 3-hydroxyglutarát a 3-hydroxybutyřát (první 3 zleva).

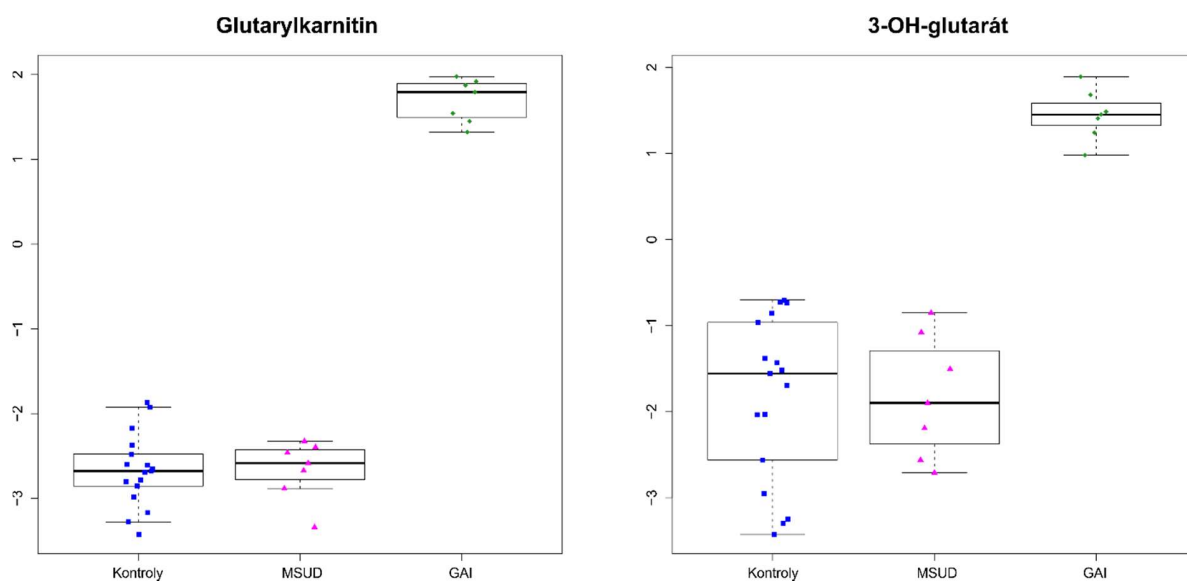
4) Ověření závěrů vícerozměrné analýzy. Tento krok je nezbytný, protože vícerozměrné statistické metody mohou při projekci veškeré proměnlivosti v datech do většinou dvou hlavních

komponent data zkreslit. Příčinou může být např. nesprávně zvolené škálování proměnných (které neodpovídá struktuře dat), na které je PCA velice citlivá [82]. V některých případech lze chybu ve statistickém vyhodnocení dat odhalit již ze samotné PCA, když např. data místo shluků vykazují lineární trend a tvoří na rozptylovém diagramu komponentních skóre podkovu (angl. „horseshoe effect“ [90,91], Obr. 11).



Obr. 11: Chybný PCA model poukazující na lineární trend v datech (vlastní data, cílená LC-MS/MS analýza krevních skvrn; P-pacienti, K-kontroly)

Ve většině případů ale bývají chyby modelu skryty. Proto je třeba závěry vícerozměrné analýzy dat ověřit na jednorozměrných datech (jednotlivých analytech). K tomu se s výhodou používají např. krabicové grafy (Obr. 12), které vizualizují data pro danou proměnnou pomocí kvartilů. Krabice grafu je dole tvořena 1. a nahoře 3. kvartilem, vnitřní část krabice protíná medián. Vousy, vycházející ze středu horní a dolní části krabice směrem nahoru a dolů, vyjadřují variabilitu dat nad třetím a pod prvním kvartilem. Mimo vousy se mohou vyskytovat odlehlé hodnoty (angl. „outliery“) jako jednotlivé body. Při ověřování závěrů vícerozměrné statistické analýzy se zobrazují krabicové grafy reprezentující distribuci abundancí daného analytu v rámci skupiny pro jednotlivé skupiny vedle sebe. U signifikantně odlišných metabolitů je na první pohled patrný rozdíl mediánů, ani interkvartilová rozpětí se nepřekrývají.



Obr. 12: Krabicové grafy porovnávající hladinu glutarylkarnitinu a 3-hydroxyglutarátu v séru zdravých kontrol, pacientů s MSUD a GAI. Oba biomarkery signifikantně odlišné pouze u GAI, zatímco pro MSUD ne. Jedná se o první 2 analyty z VIP plotu na Obr. 10, zobrazujícím porovnání GAI a kontrol.

## 7 Cíle práce

Cíle disertační práce lze shrnout do následujících bodů:

- I. Vyvinout přístup přípravy vzorku pro analýzu metabolomu pomocí GC-MS založený na urease imobilizované na magnetických mikročásticích a porovnat jej s dalšími metodami přípravy vzorku (kyselou extrakcí do ethylacetátu, volnou ureasou, vzorkem bez úpravy).
- II. Vyvinout způsob škálování 2D separačního prostoru, který by co nejlépe odrážel využitelný separační prostor a byl nezávislý na vzorku, použitém pro optimalizaci.
- III. Najít nejvhodnější způsob určení ortogonality pro nízkodimenzionální a vysokodimenzionální vzorky.
- IV. Pomocí ortogonality určit nejvhodnější kombinaci kolon pro GC×GC analýzu moči a porovnat ji s kombinací pro plasmatické vzorky.

Řešení cílů vedlo k publikačním výstupům:

Jaroslava Jáčová, Miroslav Jořenek, Kristýna Pospíšková, Lukáš Najdekr, Ludmila Zajoncová, David Friedecký, Tomáš Adam: *Urease-immobilized magnetic microparticles in urine sample preparation for metabolomic analysis by gas chromatography-mass spectrometry*, odesláno do J Chromatogr A

Jáčová J, Gardlo A, Dimandja JD, Adam T, Friedecký D: *Impact of sample dimensionality on orthogonality metrics in comprehensive two-dimensional separations*. Anal Chim Acta. 2019 Aug 8;1064:138-149. doi: 10.1016/j.aca.2019.03.018. Epub 2019 Mar 12.

Jáčová J, Gardlo A, Friedecký D, Adam T, Dimandja JD: *Sample-independent approach to normalize two-dimensional data for orthogonality evaluation using whole separation space scaling*. J Chromatogr A. 2017 Aug 18;1511:1-8. doi: 10.1016/j.chroma.2017.06.076. Epub 2017 Jul 3.

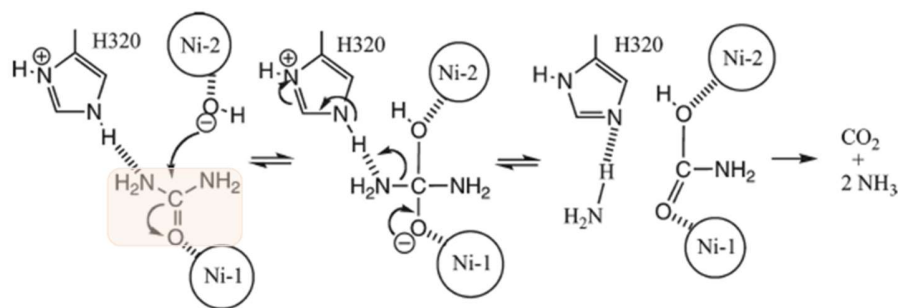
## 8 Imobilizovaná ureasa pro GC-MS metabolomiku moči

Močovina komplikuje GC-MS analýzu, protože jako konečný produkt metabolismu bílkovin je v moči přítomna ve vysoké koncentraci. Při přípravě vzorku pak dochází ke kompetici mezi analyty a močovinou při derivatizaci, a tak je řada analytů derivatizována jen částečně anebo vůbec. V separaci jsou místo jednoho derivatizačního produktu daného analytu přítomny dva nebo více, což má zásadní vliv na limit detekce těchto látek. Volné vodíkové atomy nederivatizovaných skupin se ireverzibilně váží na povrch stacionární fáze, čímž se zhoršuje tvar píku analytu i životnost kolony [92]. Samotná močovina v koloně působí jako separační interference: vyvolává přetížení kolony, její tři píky (nederivatizovaná močovina, močovina derivatizovaná dvěma a třemi trimethylsilylovými (TMS) skupinami) maskují řadu analytů a tím znesnadňují jejich detekci či kvantifikaci. Proto je nezbytné obsah močoviny před GC-MS analýzou moči snížit nebo ji zcela odstranit.

### 8.1 Metody přípravy vzorku pro analýzu metabolomu moči pomocí GC-MS

Tradičním a dosud rutinně používaným přístupem pro přípravy vzorku moči pro GC-MS analýzu je kyselá extrakce do ethylacetátu [20]. Močovina je při pH 1-2 ionizována, a tak do ethylacetátu přechází jen v omezené míře. Zároveň však dochází k ionizaci a tím i snížené extrakci celé řady dalších metabolitů moči (např. aminokyselin, hydroxykyselin, acylglycinů, acylkarnitinů, purinů, pyrimidinů). Metabolom moči se tak výrazně změní. Navíc účinnost extrakce je proměnlivá, což má za následek vysokou analytickou variabilitu.

Alternativním přístupem je aplikace ureasy [63,64,92–94], enzymu izolovaného z kanaválie mečovitě (*Canavalia ensiformis*, někdy též nazývané bob koňský nebo fazole gotani). Tento enzym ze třídy hydrolas (E.C. 3.5.1.5) rozkládá močovinu na oxid uhličitý a amoniak (Obr. 13) [95,96]. Je vysoce specifický, z dalších substrátů hydrolyzuje pouze hydroxymočovinu (lék pro léčbu klinické myeloidní leukémie [97]) na  $\text{HCO}_3^-$ ,  $\text{NH}_4^+$  a  $\text{NH}_2\text{OH}$  [96]. Tento přístup má tři hlavní nevýhody: je poměrně pracný [64,92], zanáší do analýzy řadu artefaktů (převážně při odstraňování ureasy deproteinací) [63,64,66,93] a vyvolává změny hladin některých metabolitů [63,98]. V této práci je nově využito ureasy imobilizované na magnetický nosič. Je u ní testován vliv imobilizace na další nakládání, množství artefaktů a na metabolom moči jako celek a porovnán s ostatními přístupy přípravy vzorku.



Obr. 13: Ureasou katalyzovaná hydrolýza močoviny. Aktivní místo pro navázání močoviny je tvořeno histidinem (H320), který slouží jako donor vodíkových atomů. První atom  $\text{Ni}^{2+}$  se chová jako Lewisova kyselina a aktivuje karbonylový kyslík (Ni-1), druhý atom  $\text{Ni}^{2+}$  váže vodu (Ni-2). Převzato z [96].

## 8.2 Materiál

N,O-bis(trimethylsilyl)trifluoroacetamid s 1 % trimethylchlorosilanu (BSTFA), ethylacetát, O-ethylhydroxylamin hydrochlorid, pyridin, 4-fenylbutyrát (interní standard, IS), standardní směs alkanů pro testování GC systémů (standard  $\text{C}_{10}\text{-C}_{40}$ ), ureasa (prášková, typ IX, 50,000-100,000 jednotek/g pevné látky, označení U4002), mevalonát sodný, mevalonolakton, N-acetylaspartát, a rozpouštědla pro LC-MS (voda, methanol, acetonitril a hydroxid amonný) jsou zakoupeny od Sigma-Aldrich (St. Louis, Missouri, USA), chlorotrimethylsilan (TMCS) od Merck KGaA (Darmstadt, Spolková republika Německo), bezvodý síran sodný, chlorid sodný a kyselina chlorovodíková od mikroCHEM (Pezinok, Slovenská republika), hexanoylglycin od VU Medical Center (Amsterdam, Nizozemské království) a kontrolní standardní vzorek Control Organic Acids (COA) od MCA Laboratory (Winterswijk, Nizozemské království, číslo šarže 2013.007), všechny alespoň v analytické čistotě. Ureasa imobilizovaná na magnetické mikročástice byla vyvinuta na Katedře biochemie Přírodovědecké fakulty Univerzity Palackého v Olomouci [99].

## 8.3 Vzorky

V experimentu je použito dvou typů vzorků: kontrolního materiálu na bázi lidské moči (COA, z angl. „control organic acids“), do které byla výrobcem přidána známá koncentrace 23 metabolitů (Tab. 1), a reálných močí pacientů s glutarovou acidurií typu I a zdravých kontrol. COA je skladován v lednici ( $4\text{ }^\circ\text{C}$ ), před použitím rekonstituován 10 ml vody pro LC-MS, uzavřen zátkou a ponechán 15 min při pokojové teplotě. Poté je 20 min promícháván na rolovací třepačce. Reálné vzorky moči jsou co

nejdříve po odběru zamražený a skladovány při -80 °C. Pro analýzu je u obou typů vzorků použit objem odpovídající koncentraci kreatininu 1 mmol.l<sup>-1</sup>.

Tab. 1: Analyty obsažené v COA a jejich koncentrace stanovené rutinními uživateli. Koncentrace kreatininu je 4,04 mmol.l<sup>-1</sup>.

Analyt	Konsensuální hodnota [μmol.l <sup>-1</sup> ]	95% konfidenční interval	Očekávané rozmezí laboratorních výsledků
2-hydroxyglutarát	185	169 – 201	115 – 225
3-methylglutarát	66	63 – 69	54 – 78
3-hydroxy-3-methylglutarát	143	107 – 179	5 – 281
3-hydroxyisobutyrate	103	76 – 130	3 – 203
3-hydroxyisovalerát	188	168 – 208	101 – 275
4-hydroxybutyrát	83,4	67 – 100	16 – 151
Adipát	176	165 – 187	126 – 226
Glycerát	548	474 – 622	252 – 844
Ethylmalonát	58	54 – 62	40 – 76
Fumarát	61,3	57 – 66	42 – 80
Glutarát	104	98 – 110	75 – 133
Glykolát	192	177 – 207	123 – 261
Hexanoylglycin	36,2	33 – 40	23 – 50
2-oxoglutarát	207	189 – 225	128 – 286
Malát	118	101 – 135	59 – 177
Methylmalonát	794	741 – 847	529 – 1059
Mevalonát	397	351 – 443	225 – 569
N-acetylaspartát	635	525 – 745	175 – 1095
Pyroglutamát	494	415 – 573	159 – 829
Sebakát	176	161 – 191	108 – 244
Suberát	146	137 – 155	104 – 188
Tiglylglycin	86,6	77 – 96	51 – 122
Vanillaktát	39,3	31 – 48	16 – 62

### 8.3.1 Příprava vzorků COA pro GC-MS analýzu

Vzorek COA je pro GC-MS analýzu připraven 4 různými metodami (vždy 6 replikátů + blank pro každou metodu):

- 1) Kyselá extrakce do ethylacetátu (EX): 495  $\mu\text{l}$  COA je napipetováno do zábrusové skleněné zkumavky obsahující 20  $\mu\text{l}$  IS (2,11  $\text{mmol.l}^{-1}$  v methanolu), poté je přidána malá lžička O-ethylhydroxylamin hydrochloridu (přibližně 0,03 g) a objem vzorku doplněn vodou do 2 ml. Po 30 min stání při pokojové teplotě jsou přidány 3 kapky HCl (1:1) a jedna lžička NaCl (1,0 g). Extrakce do 4 ml ethylacetátu je provedena třikrát: po 10 min třepání na třepačce je odebrána horní vrstva a přenesena do zkumavky. K 12 ml extraktu ve zkumavce je přidána lžička bezvodého  $\text{Na}_2\text{SO}_4$  (1,0 g), obsah zkumavky dobře promíchán. Po 10 min stání při pokojové teplotě je roztok dekantován do čisté zkumavky a odfoukán proudem dusíku při 30 °C. Odparek je 2x vymyt 400  $\mu\text{l}$  methanolu, a pak ještě jednou 300  $\mu\text{l}$  acetonu do krimpovací vialky pro GC-MS, znovu odfoukán proudem dusíku při 30 °C. Nakonec je přidáno 300  $\mu\text{l}$  derivatizační směsi (BSTFA:pyridin:TMCS, 25:25:1 (v/v/v)), vialka je uzavřena pod dusíkem, její obsah promíchán a 20 min derivatizován při 70 °C.
- 2) Volná ureasa (FUT): 495  $\mu\text{l}$  COA je smícháno s 60  $\mu\text{l}$  ureasy (0,6  $\text{mg.ml}^{-1}$  ve fyziologickém roztoku) a 445  $\mu\text{l}$  vody, 40 min promícháváno pomocí míchačky při 37 °C (72 g). 500  $\mu\text{l}$  reakční směsi je přeneseno do čisté mikrozkušavky, poté je přidán 1 ml ledového methanolu, vzorek je vložen do mrazáku (-80 °C), po 30 min centrifugován (14 000 g; 8 min). Supernatant je přenesen do skleněné vialky obsahující 10  $\mu\text{l}$  IS (2,11  $\text{mmol.l}^{-1}$  v methanolu), vzorek je opět vložen do mrazáku (-80 °C), po 30 min je přenesen do lyofilizátoru a zlyofilizován (obvykle přes noc). Následně je vzorek derivatizován ve dvou krocích: ethoximace (10  $\mu\text{l}$  ethylhydroxylamin hydrochloridu v pyridinu o koncentraci 56  $\text{mg.ml}^{-1}$  a 140  $\mu\text{l}$  pyridinu) po 60 min při 50 °C, trimethylsilylace (150  $\mu\text{l}$  silylační směsi BSTFA+TMCS:TMCS, 25:1) po 20 min při 70 °C.
- 3) Imobilizovaná ureasa (IUT): 495  $\mu\text{l}$  COA je smícháno s 100 mg mikroparticulí s imobilizovanou ureasou a 1505  $\mu\text{l}$  vody, 40 min mícháno na míchačce při 37 °C (72 g). Magnetické mikročástice jsou potom odděleny pomocí magnetu. 1 ml reakční směsi je přeneseno do vialky obsahující 10  $\mu\text{l}$  IS (2,11  $\text{mmol.l}^{-1}$  v methanolu), vloženo do mrazáku (-80 °C), po 30 min jsou přeneseno do lyofilizátoru a zlyofilizováno (obvykle přes noc). Následně je vzorek derivatizován ve dvou krocích: ethoximace (10  $\mu\text{l}$  ethylhydroxylamin hydrochloridu v pyridinu o koncentraci 56  $\text{mg.ml}^{-1}$  a 140  $\mu\text{l}$  pyridinu) po 60 min při 50 °C, trimethylsilylace (150  $\mu\text{l}$  silylační směsi BSTFA+TMCS:TMCS, 25:1) po 20 min při 70 °C.



- 4) Vzorky bez úpravy (NT): 248  $\mu\text{l}$  COA je přidáno do vialky s 10  $\mu\text{l}$  IS (2,11  $\text{mmol.l}^{-1}$  v methanolu), poté je vzorek vložen do mrazáku ( $-80\text{ }^{\circ}\text{C}$ ), po 30 min přenesen do lyofilizátoru a lyofilizován (obvykle přes noc). Následně je vzorek derivatizován ve dvou krocích: ethoximace (10  $\mu\text{l}$  ethylhydroxylamin hydrochloridu v pyridinu o koncentraci 56  $\text{mg.ml}^{-1}$  a 140  $\mu\text{l}$  pyridinu) po 60 min při  $50\text{ }^{\circ}\text{C}$ , trimethylsilylace (150  $\mu\text{l}$  silylační směsi BSTFA+TMCS:TMCS, 25:1) po 20 min při  $70\text{ }^{\circ}\text{C}$ .

### 8.3.2 Příprava vzorků COA pro LC-MS/MS analýzu

Pro LC-MS/MS analýzu jsou vzorky (vždy 6 replikátů a blank) připraveny jen pomocí FUT a IUT a porovnány se vzorky ponechanými bez úpravy. EX nelze zařadit do srovnání, protože tento protokol vyžaduje ethoximaci na samém počátku přípravy vzorku. Ethoximované deriváty mají jiné MRM přechody než původní analyty, a tak by nebylo možné použít stávající metodu pro cílenou metabolomiku. Pokud by se ethoximace při přípravě vzorku vynechala, mohlo by dojít ke změně extrakční účinnosti. Postup přípravy vzorků je následující:

- 1) Volná ureasa (FUT): 495  $\mu\text{l}$  COA je smícháno s 60  $\mu\text{l}$  ureasy (0,6  $\text{mg/ml}$  ve fyziologickém roztoku) a 445  $\mu\text{l}$  vody, 40 min mícháno na míchačce při  $37\text{ }^{\circ}\text{C}$  (72 g). 500  $\mu\text{l}$  reakční směsi je přeneseno do čisté mikrozkušavky, poté je přidán 1 ml ledového methanolu a mikrozkušavka vložena do mrazáku ( $-80\text{ }^{\circ}\text{C}$ ). Po 30 min je vzorek centrifugován (14 000 g; 8 min). Supernatant (100  $\mu\text{l}$ ) je přenesen do čisté skleněné vialky.
- 2) Imobilizovaná ureasa (IUT): 495  $\mu\text{l}$  COA je smícháno se 100 mg mikroparticulí s imobilizovanou ureasou a 1505  $\mu\text{l}$  vody, 40 min mícháno na míchačce při  $37\text{ }^{\circ}\text{C}$  (72 g). Magnetické mikročástice jsou odděleny pomocí magnetu, 100  $\mu\text{l}$  reakční směsi je přeneseno do čisté skleněné vialky.
- 3) Vzorky bez úpravy (NT): 495  $\mu\text{l}$  COA je smícháno s 1505  $\mu\text{l}$  vody, 40 min mícháno na míchačce při  $37\text{ }^{\circ}\text{C}$  (72 g). 100  $\mu\text{l}$  reakční směsi je přeneseno do čisté skleněné vialky.

### 8.3.3 Příprava vzorků pacientů

10 vzorků pacientů s glutarovou acidurií typ I, 10 vzorků kontrol a jeden blank je připraveno pomocí IUT. QC je připraven smícháním 10  $\mu\text{l}$  z každého vzorku (vyjma blanku). Vzorky jsou analyzovány pomocí GC-MS, a to v náhodném pořadí. QC je zařazeno jako každý 6. vzorek.

## 8.4 Metody

### 8.4.1 GC-MS analýza a procesování dat

Vzorky pro GC-MS jsou analyzovány na plynovém chromatografu spojeném s TOF MS (Pegasus4, LECO, St. Joseph, Michigan, USA), vybaveném kolonou ZB-5MS (30 m; 0,25 mm; 0,25  $\mu$ m; Phenomenex, Torrance, California, USA). Jako nosný plyn je použito helium v režimu konstantního průtoku (1 ml.min<sup>-1</sup>). Podíl vzorku (1  $\mu$ l) je zaváděn do proudu nosného plynu metodou horké jehly (angl. "hot needle" [100]) při 250 °C, v režimu s děličem toku s dělicím poměrem 1:10. Je využito analýzy s teplotním programem, kdy je teplota prvních 5 min analýzy udržována na 80 °C, pak je rychlostí 5 °C.min<sup>-1</sup> zvyšována až na 320 °C a po posledních 10 min je udržována konstantní. Do hmotnostního spektrometru jsou analyty převáděny při 250°C. Při stejné teplotě pracuje i iontový zdroj. Jedná se o elektronovou ionizaci (EI) s urychlujícím potenciálem 70 eV. Data jsou sbírána po uplynutí úvodních 250 s analýzy, a to v rozsahu 50-500  $m/z$  s frekvencí 5 Hz. Napětí na detektoru je nastaveno na -1797 V.

Retenční indexy jsou počítány pomocí směsi alkanů (C<sub>10</sub>-C<sub>36</sub>; 5  $\mu$ g/ml v derivatizační směsi, měřeny jako samostatný vzorek), s využitím tzv. metody pro retenční indexy (Retention Index Method, ChromaTOF software, LECO).

K vyhledávání analytů pro cílenou metabolomiku je použit software s názvem „Reference“ (ChromaTOF), a to s tímto nastavením: povolená odchylka retenčního času 1 s; S/N 5; minimální shoda s NIST 700; nastavení integrace: plocha píku v dTIC, interní standard 4-fenylbutyrát, tolerance 20 %. Pro každý ze vzorků je zintegrována plocha všech 25 píků manuálně zkontrolována, exportována do MS Excel 2016 (Microsoft, Redmond, USA), normalizována plochou interního standardu. Plochy píků dvou derivatizačních produktů kyseliny mevalonové (mevalonolakton-1 TMS a 3,5-dihydroxy-3-methylpentanová kyselina-2 TMS) jsou před statistickým zpracováním sečteny. Data ve finálním datovém souboru jsou škálována na medián každé skupiny. Datový soubor je procesován v R-software (verze 3.3.0, <https://www.r-project.org>) pomocí domácího softwarového balíčku, za použití se *clr* transformace. Ke zobrazení dat je použito PCA, PLS-DA a krabicových grafů.

Pro necílenou GC-MS analýzu je použito stejných surových dat (z angl. „raw data“) jako pro cílenou analýzu. Protože jsou metody přípravy vzorku výrazně odlišné, výjimečně nejsou připraveny QC, porovnatelnost analýz je zajištěna výhradně pomocí IS. Data jsou z ChromaTOF software vyexportována ve formátu Andi MS (.cdf) a vložena do online verze XCMS software (verze 2.2.3) [101,102], s následujícími parametry: detekce potenciálních metabolitů: *centWave* (300 ppm; S/N 25; peak width 5-10 s; mzdiff 0,1; prefilter 1), *peaks* (3; prefilter intensity 100; noise 0), *grouping* (method density; bw 10; mz width 0,25; min frac 0,5; minsamp 1). Pomocí výše uvedeného algoritmu jsou na

základě zadaných parametrů jsou v chromatogramech vyhledány píky potenciálních metabolitů, které odpovídají kritériím. Tyto potenciální metabolity, anglicky označované jako „features“, jsou jako konečný výstup uvedeny v tabulce výsledků XCMS. Dále jsou upravovány takto: potenciální metabolity s  $m/z$  147 a pod 100 jsou odstraněny (fragment 147 náleží silylačnímu činidlu, fragmenty menší než 100 nejsou specifické). Jelikož se jedná o EI spektra, každý z píků analytů je zastoupen hned několika fragmenty (resp. features). Tyto fragmenty jsou pomocí XCMS algoritmu sloučeny do skupin (každá skupina odpovídá jednomu píku). Z každé skupiny je do finálního datového souboru za vybrán jen jeden fragment, který má nejvyšší intenzitu. Datový soubor je procesován v R-software (verze 3.3.0, <https://www.r-project.org>) pomocí softwarového balíčku vyvinutého v naší laboratoři [103], za použití *clr* transformace. Ke zobrazení dat slouží PCA, PLS-DA a krabicové grafy.

#### 8.4.2 LC-MS analýza a procesování dat

Pomocí kapalinového chromatografu UltiMate 3000 RS (Dionex, Sunnyvale, CA, USA) spojeného s trojitým kvadrupólem Triple Quad 6500 (SCIEX, Framingham, MA, USA) je za použití metody převzaté od Bajad a kol. [104] ve vzorku sledován obsah 346 analytů. Pro separaci na principu hydrofilní interakční chromatografie je použita kolona Luna NH<sub>2</sub> (3.0  $\mu$ m, 2 x 100 mm, Phenomenex, Torrance, USA), chráněná předkolonou vyrobenou ze stejného materiálu (4 x 2 mm), a vyhřívána na 35 °C. Mobilní fáze A je tvořena z vodného roztoku octanu amonného (20 mmol.l<sup>-1</sup>, pH 9,75), mobilní fáze B čistým acetonitrilem. Je využito následující gradientové eluce: 0-7 min: 95% → 10% B; 7-13 min: 10% B; 13-13,5 min: 10% → 95% B; 13,5-7 min: 95% B, průtoková rychlost 0,3 ml.min<sup>-1</sup>.

Odezva metabolitů je po ionizaci elektrosprejem měřena v režimu sledování více vybraných iontových reakcí v definovaných časových oknech s definovanou váhou pro jednotlivé metabolity (angl. „weighted scheduled MRM“), a to současně v obou módech, za režimu přepínání polarit (angl. „polarity switching“). Rozlišení obou kvadrupólů hmotnostního analyzátoru je jednotkové. Parametry iontového zdroje jsou nastaveny takto: napětí na sprejovací kapiláře je +5500/-4500 V, teplota iontového zdroje 400 °C, tlaky obou zmlžujících plynů 40 psi, clonícího plynu 30 psi. Vysoce čistý dusík je použit jako kolizní plyn v kolizní cele i jako pomocný plyn v iontovém zdroji. Parametry specifické pro každou látku (deklasterační potenciál, vstupní potenciál, kolizní energie, výstupní potenciál) jsou optimalizovány pomocí standardů v Analyst software (verze 1.6.2 SCIEX). Analyty jsou dle retenčního času a specifických MRM přechodů detekovány a identifikovány v MultiQuant software (verze 3.0, SCIEX), kde jsou také integrovány plochy píků.

Datový soubor je korigován a normalizován pomocí LOESS. Metabolity, jejichž relativní směrodatná odchylka je v QC vyšší než 30 %, jsou z datového souboru vyloučeny. Chybějící hodnoty jsou nahrazeny dvěma třetinami minimální hodnoty získané pro daný analyt. Datový soubor je pak pomocí *clr* transformace matematicky převeden na kompoziční data a normalizován. Analýza dat (korelační analýza, PCA a PLS-DA) je provedena pomocí MS Excel a R-software (v. 3.3.0, <https://www.r-project.org>).

## 8.5 Výsledky a diskuse

Pro vyhodnocení způsobu přípravy moči založeném na urease imobilizované na magnetických mikročásticích je aplikován komplexní metabolický přístup, při kterém se k analýze metabolomu využívá kombinace více analytických technik (nejčastěji GC-MS a LC-MS). Tento přístup, založený na předpokladu, že kombinace LC-MS a GC-MS umožňuje analýzu hydrofilních i lipofilních látek, je použit také v práci Kind a kol. [98] pro vyhodnocení protokolu založeného na volné urease.

Protože však je močovina hlavní interferující látkou GC-MS analýz, důraz je kladen zejména na vliv metody na GC-MS analýzu. Proto je v případě GC-MS pro porovnání vlivu přípravy vzorku moči pomocí IUT, FUT, EX a NT na metabolom COA použit jak cílený, tak necílený metabolický přístup, zatímco u LC-MS pouze cílený. Je však nutné zmínit, že NT není pro GC-MS metabolomiku používán a je zde zařazen jen pro srovnání. Využitelnost IUT pro metabolický experiment je ověřena pomocí močí pacientů s GAI a zdravých kontrol.

### 8.5.1 Pracnost metod

Testované metody přípravy vzorku se liší v pracnosti i časových nárocích. Nejpracnější ze všech je EX, který se skládá z 11 různých kroků a 6 h nepřetržité práce. Vzorek je však k analýze připraven týž den, protože nevyžaduje lyofilizaci. Většinou je ale obtížné připravit více než 12 vzorků najednou.

Při použití FUT lze vzorek připravit v 8 různých krocích, vyžaduje 3 h práce a lyofilizaci přes noc. Oproti tomu IUT zahrnuje pouze z 5 kroků. Protože nepotřebuje deproteinaci, k přípravě jsou nutné pouze 2 h práce a lyofilizace přes noc. Oba ureasové protokoly umožňují přípravu většího množství vzorků. Jediným omezením je kapacita rotátoru.

### 8.5.2 Kvalitativní změny metabolomu

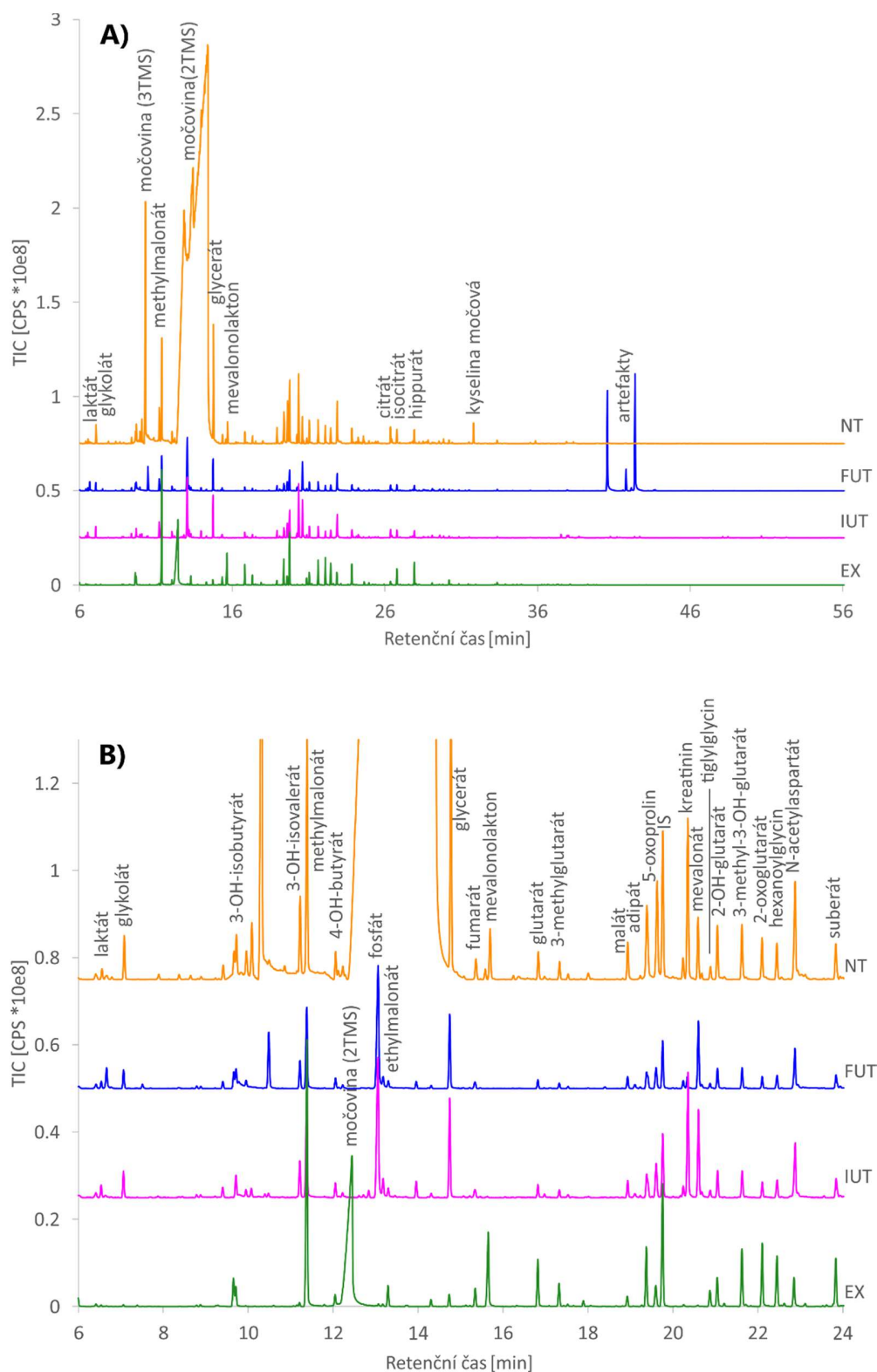
Chromatografický vzor GC-MS analýzy IUT a FUT je obdobný (Obr. 14). Hlavním rozdílem mezi analýzami je řada unikátních píků na konci analýz FUT. Tyto píky jsou přítomny i v analýzách blanků

(Obr. 15), proto se lze domnívat, že se jedná o artefakty (viz kapitola 8.5.3). Oproti tomu u EX je přítomno nejméně píků, navíc s naprosto odlišnou odezvou. NT vzorky mají nejvyšší odezvy v celkovém iontovém proudu (TIC) vzorků, obzvláště pak v retenčním okně močoviny (mezi 12. a 15. min). Ohromný pík močoviny také způsobuje přetížení kolony, a proto i posun retenčních časů látek, které jsou v koloně zadržovány více než močovina (např. glycerát, fumarát, mevalonolakton). Močovina také zhoršuje identifikaci koeluuujících látek (např. pro ethylmalonát v NT vzorcích byla shoda s NIST pouze okolo 300, což není pro identifikaci látky dostačující).

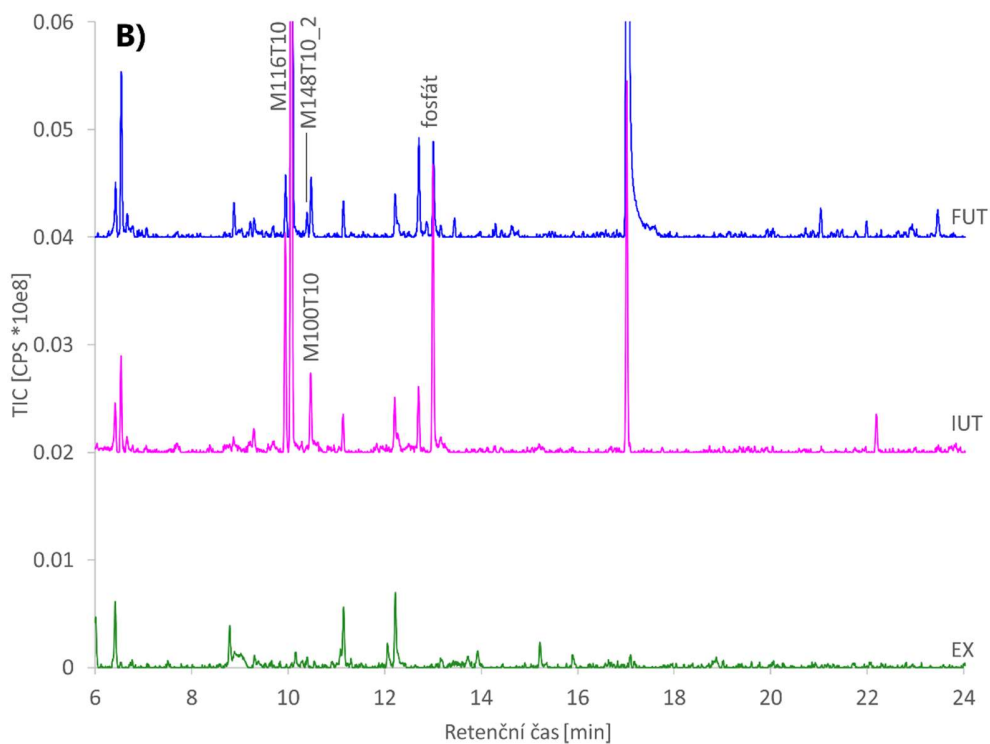
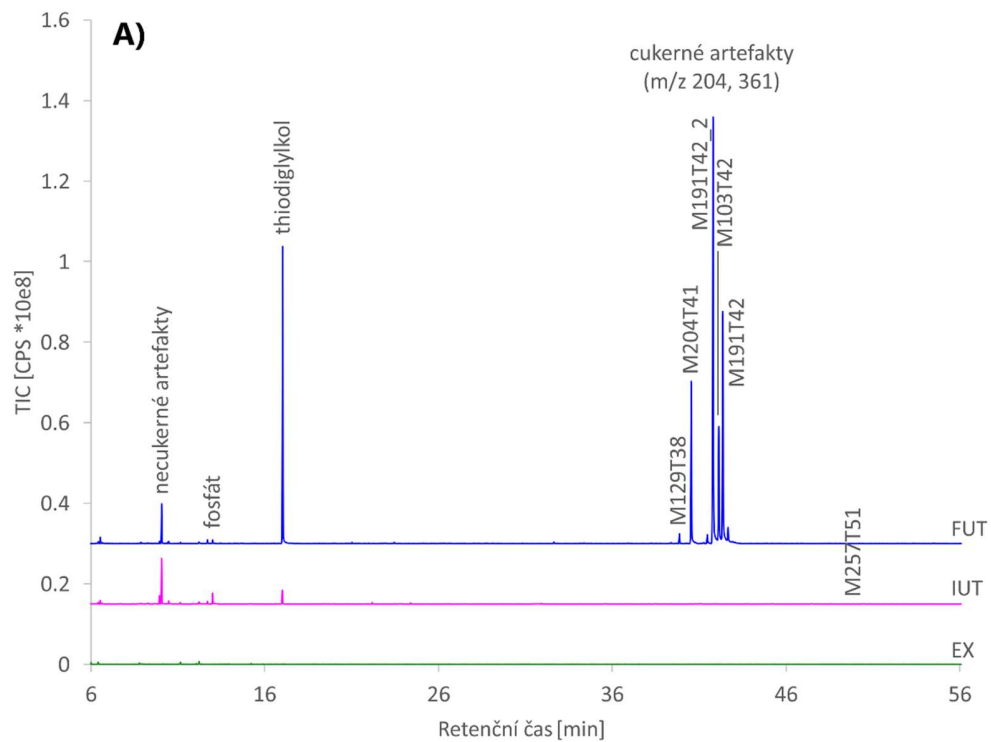
Stejný trend je patrný také při analýze vícerozměrnými statistickými metodami. PCA necílené GC-MS metabolomiky (Obr. 16) ukazuje, že vzorky připravené pomocí FUT a IUT jsou si ve srovnání s ostatními typy přípravy podobné, protože jejich shluky jsou lokalizovány blízko u sebe, ale nepřekrývají se. Vzorky připravené pomocí EX jsou na základě první významnější komponenty (PC1 59,88 %) podobné FUT/ IUT a všechny tři se výrazně liší od NT. Na bázi druhé komponenty (PC2 28,73 %) je však zřejmé, že EX se od FUT a IUT odlišuje.

Cílená GC-MS metabolomika (Obr. 17) porovnává intenzity ploch detekovaných metabolitů, které byly součástí analytické metody. Vzorky FUT a IUT tvoří na PCA kompaktní, vzájemně se překrývající shluky, na základě čehož lze soudit, že analyty mají obdobné intenzity a jejich analytická variabilita je nízká. Na základě první hlavní komponenty (PC1 84,17 %) se nejvíce odlišují od EX. Disperze vzorků EX podél druhé komponenty (PC2 7,03 %) navíc poukazuje na vysokou analytickou variabilitu extrakce, která je dokonce vyšší než mezi vzorky FUT/IUT a NT.

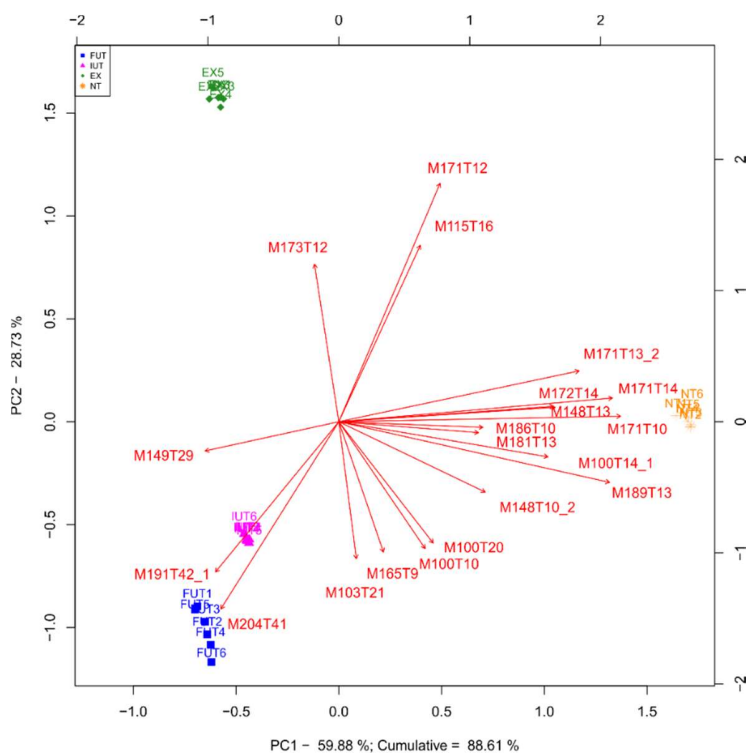
V LC-MS analýze je intenzita i variabilita metabolitů napříč jednotlivými přístupy srovnatelná (Obr. 18). Seznam všech analytů nalezených ve vzorcích technikou LC-MS je uveden v Tab. S1. PCA ukazuje (Obr. 19), že rozdílnost mezi skupinami NT a IUT je malá (částečně se dělí na základě PC2 s nízkou významností 9,23 %), zatímco FUT se od nich odlišuje (PC1 58,58 %).



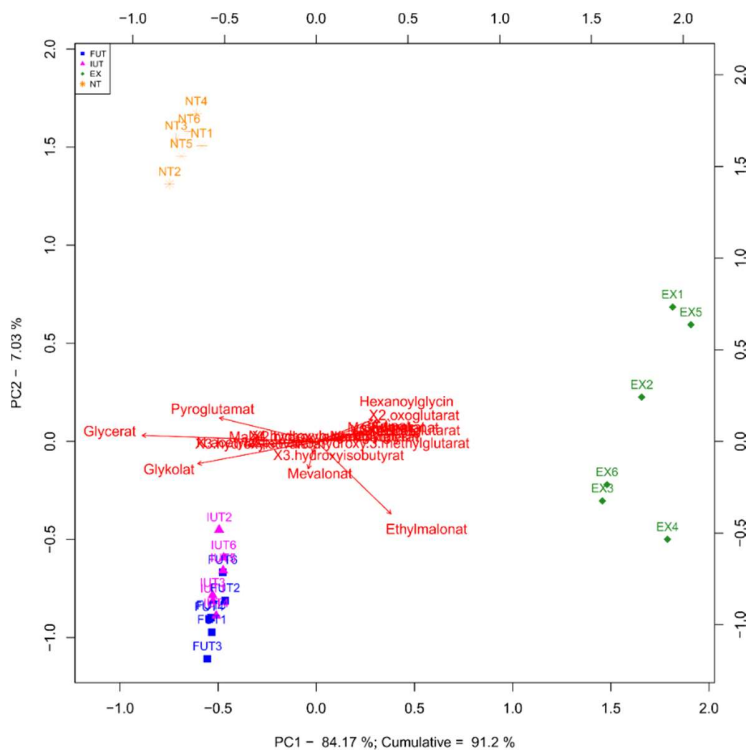
Obr. 14: GC-MS analýza COA připravených pomocí jednotlivých způsobů přípravy vzorku: A) celkový chromatografický vzor, B) detail.



Obr. 15: GC-MS analýza blanků připravených pomocí jednotlivých způsobů přípravy vzorku: A) celkový chromatografický vzor, B) detail.

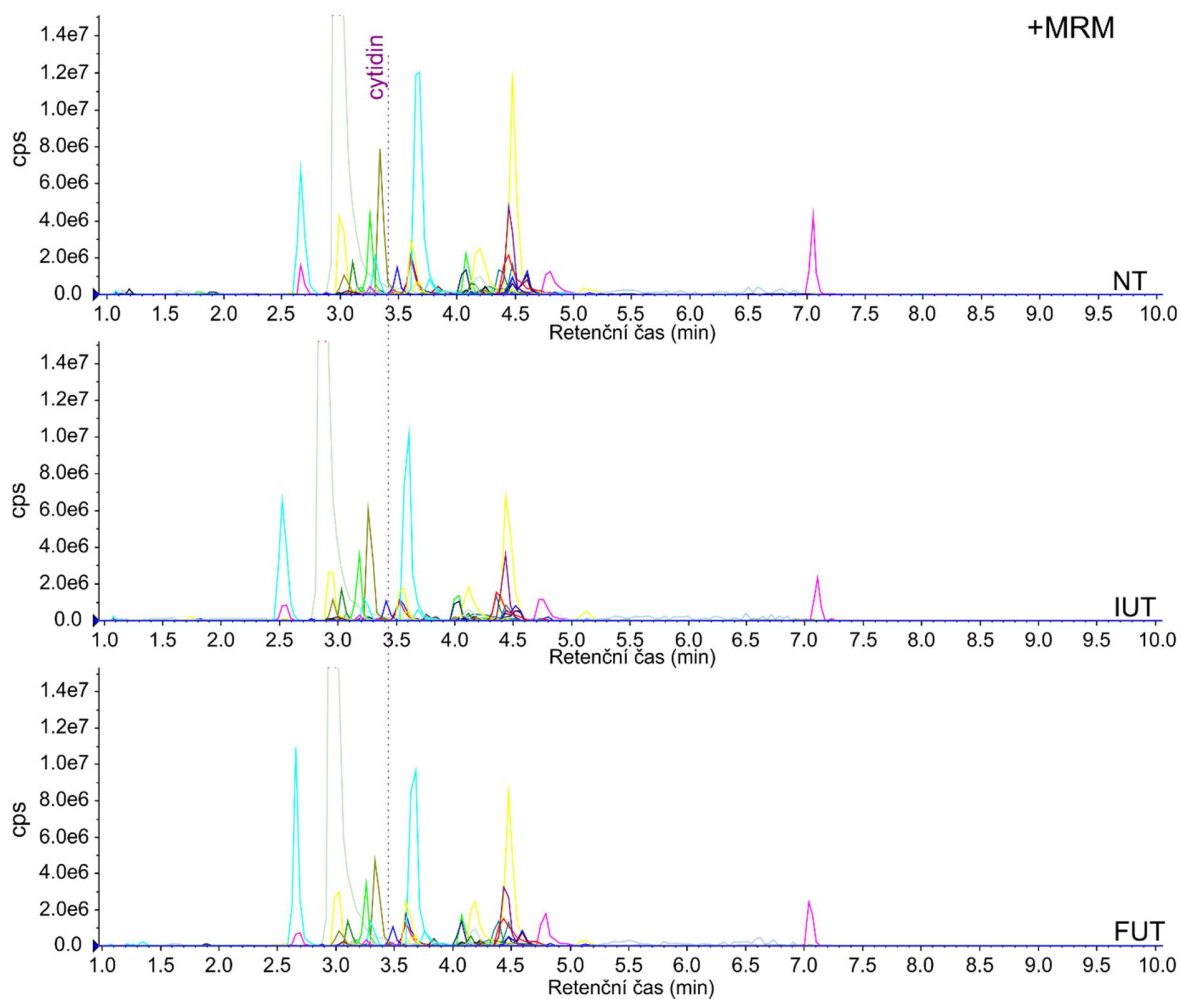


Obr. 16: Dvojný graf PCA necílené GC-MS metabolomiky

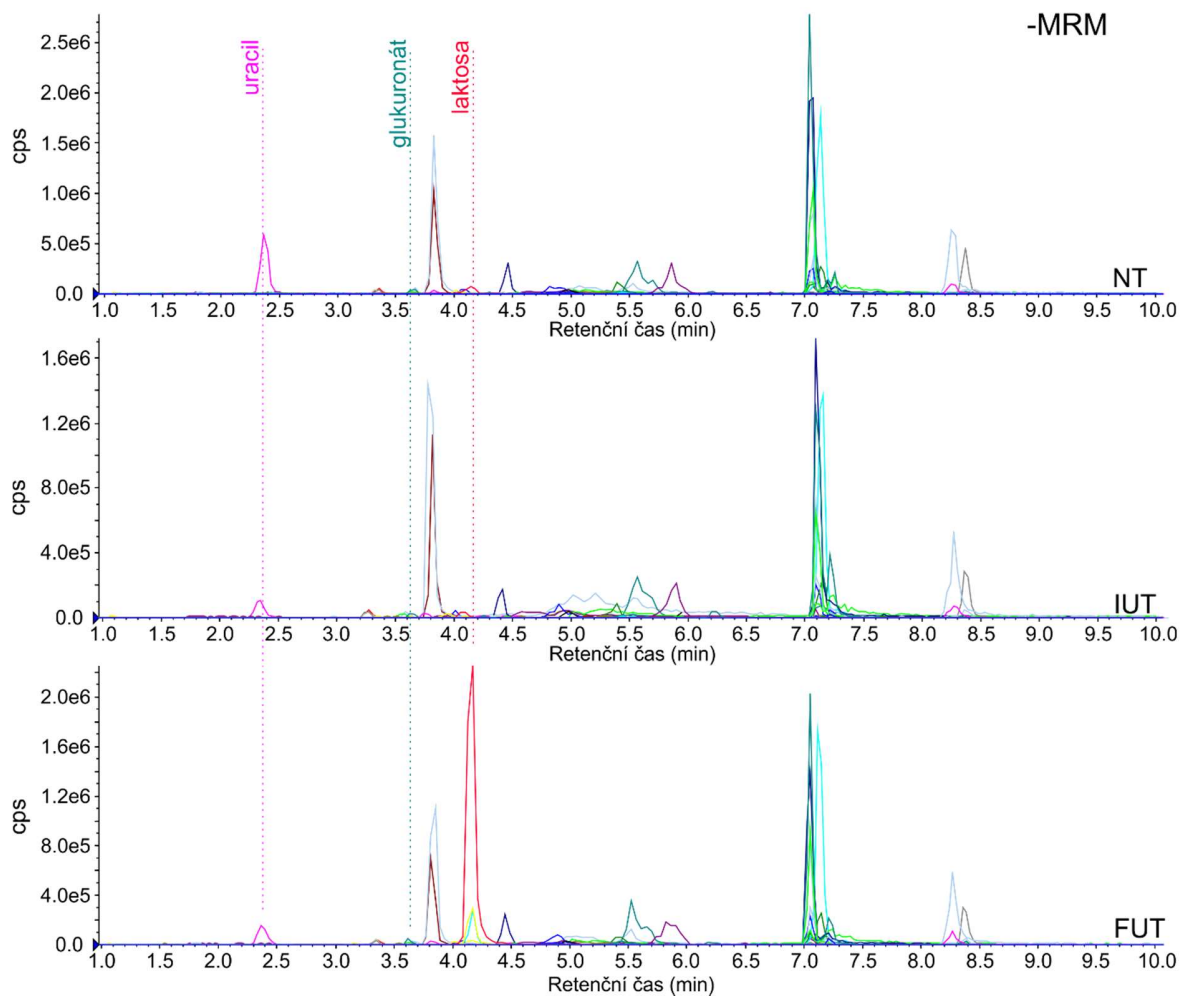


Obr. 17: Dvojný graf PCA cílené GC-MS metabolomiky

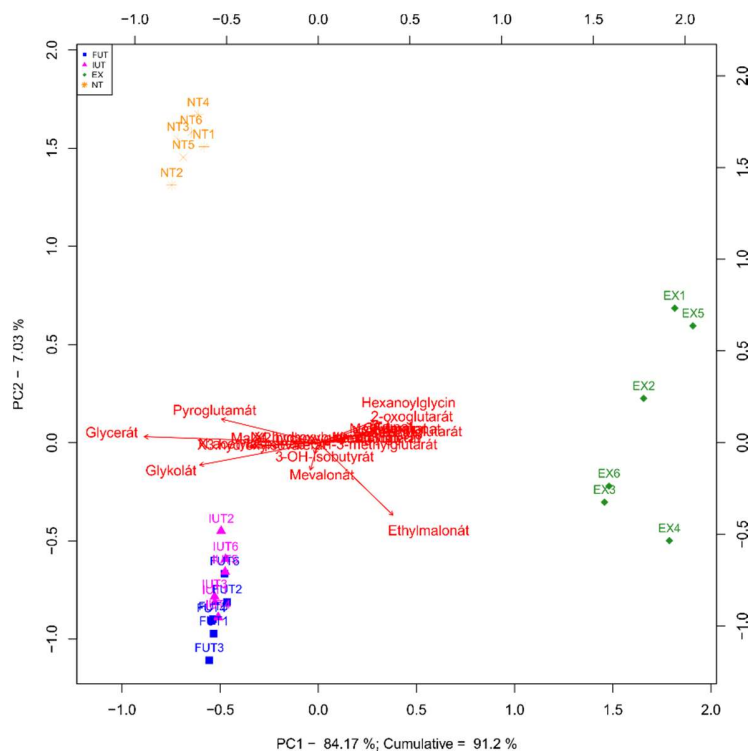




Obr. 18 A): Chromatografický vzor LC-MS, pozitivní mód. V chromatogramu je vyznačen diskriminující metabolit. Retenční časy píků ostatních analytů jsou uvedeny v Tab. S1.



Obr. 18 B): Chromatografický vzor LC-MS, negativní mód. V chromatogramech jsou vyznačeny metabolity diskriminující jednotlivé skupiny. Retenční časy píků ostatních analytů jsou uvedeny v Tab. S1.



Obr. 19: Dvojný graf PCA cílené LC-MS analýzy COA. Diskriminující metabolity: C023 – laktosa, C048 – cytidin, C032 – glukuronát, C017 – uracil. Seznam všech analytů je uveden v Tab. S1.

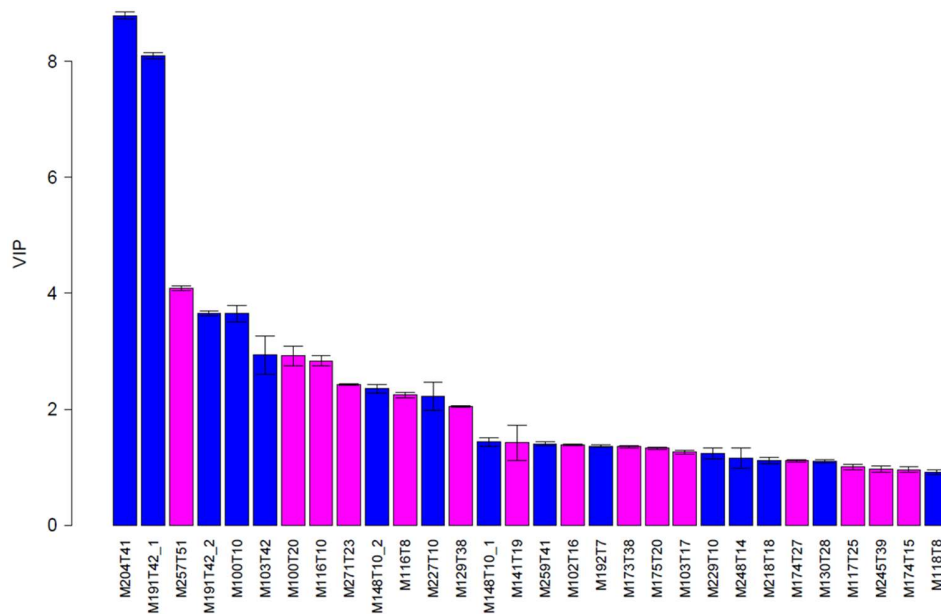
### 8.5.3 Artefakty

Jako artefakty se označují látky, které jsou přítomny v chromatografických separacích, ale nepocházejí z původního vzorku [105]. Mohou vznikat při derivatizaci (např. neobvyklé deriváty či vedlejší produkty), během přípravy vzorku (zbytky činidel, nečistoty aj.) nebo mohou být do záznamů vneseny chromatografickým systémem (např. důsledkem krvácivosti kolony či reakcí probíhajících v injektoru či iontovém zdroji) [98]. Také ureasa je známá tím, že do analýzy vnáší řadu artefaktů [63], proto je nezbytné ověřit, zda přítomnost nosiče, na který je ureasa navázána, jejich množství ještě nezvyšuje. Využívá se k tomu dvou různých experimentů: porovnání analýz blanků a necílený GC-MS metabolomický přístup.

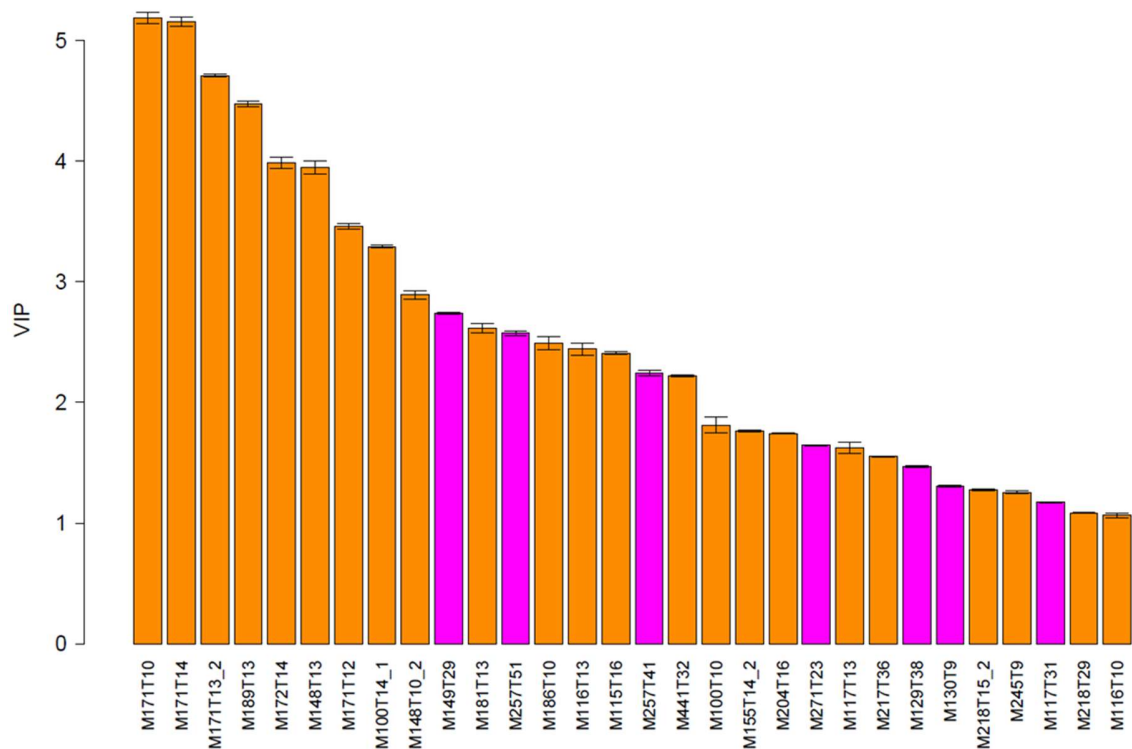
Porovnání analýz blanků ukazuje, že počet píků nalezených v separacích je pro jednotlivé postupy přípravy odlišné. U FUT je při S/N 1000 přítomno 18 píků, u IUT 9 a u EX pouze 4 (Obr. 15). Z necílené GC-MS metabolomiky vyplývá, že artefakty M204T41, M191T42\_1, M191T42\_2, M257T51, M129T38, and M103T42 tvoří většinu diskriminujících látek mezi IUT a FUT vzorky (Obr. 20). Zatímco

u IUT a NT (Obr. 21) jsou to především deriváty močoviny (M171T10, M171T12, M171T14) a látky, které mají podobnou retenci jako ona (ethylmalonát - M171T13\_2, fosfát - M189T13, sukcinát - M172T14, glycerol - M148T13, serin - M116T13), u EX a IUT (Obr. 22) je to močovina (171T12), kreatinin (M100T20) a další metabolity, ovlivněné účinností extrakce. Signifikantně odlišným metabolitem mezi NT a IUT i EX a IUT je mevalonát, protože tato látka se v nízkém pH (u EX) vyskytuje ve formě cyklického vnitřního esteru, tzv. mevalonolaktonu (M115T16), ve vysokém pH (FUT a IUT) v lineární podobě (M103T21), v neutrálním pH NT vzorku jsou přítomny obě formy.

Nejvýraznějšími artefakty v separaci FUT jsou tvořeny skupinou píků na konci analýzy ( $t_r$  vyšší než 37 min, Obr. 14). Protože ve svých spektrech obsahují fragment s  $m/z$  204, je zřejmé, že se jedná o některý z cukrů, pravděpodobně disacharid ( $m/z$  361, vysoký retenční čas). Protože EI spektra sacharidů jsou si velmi podobná, není možné tento disacharid identifikovat pomocí NIST ani GOLM databáze. I na LC-MS má laktosa se sacharosou a dokonce i s maltosou a melibiosou stejný MRM přechod. Ani retenční časy a indexy nejsou příliš nápomocné, protože retence laktosy a sacharosy je u obou technik obdobná. Řada studií však uvádí, že po aplikaci FUT se ve vzorku zvýšil obsah laktosy, která se používá ke stabilizaci ureasy [64,106,107]. Při aplikaci IUT se obsah těchto artefaktů výrazně snížil, pravděpodobně důsledkem řady promývacích kroků, které jsou provedeny během imobilizace. V GC-MS separacích IUT tyto píky zcela chybí, v citlivější LC-MS analýze jsou přítomna stopová množství, a to dokonce i v NT vzorcích (v tomto případě se zřejmě jedná o endogenní metabolit). Intenzita „laktosy“ byla ve vzorcích FUT 57× a ve vzorcích IUT 2× vyšší než v NT (Tab. S1).



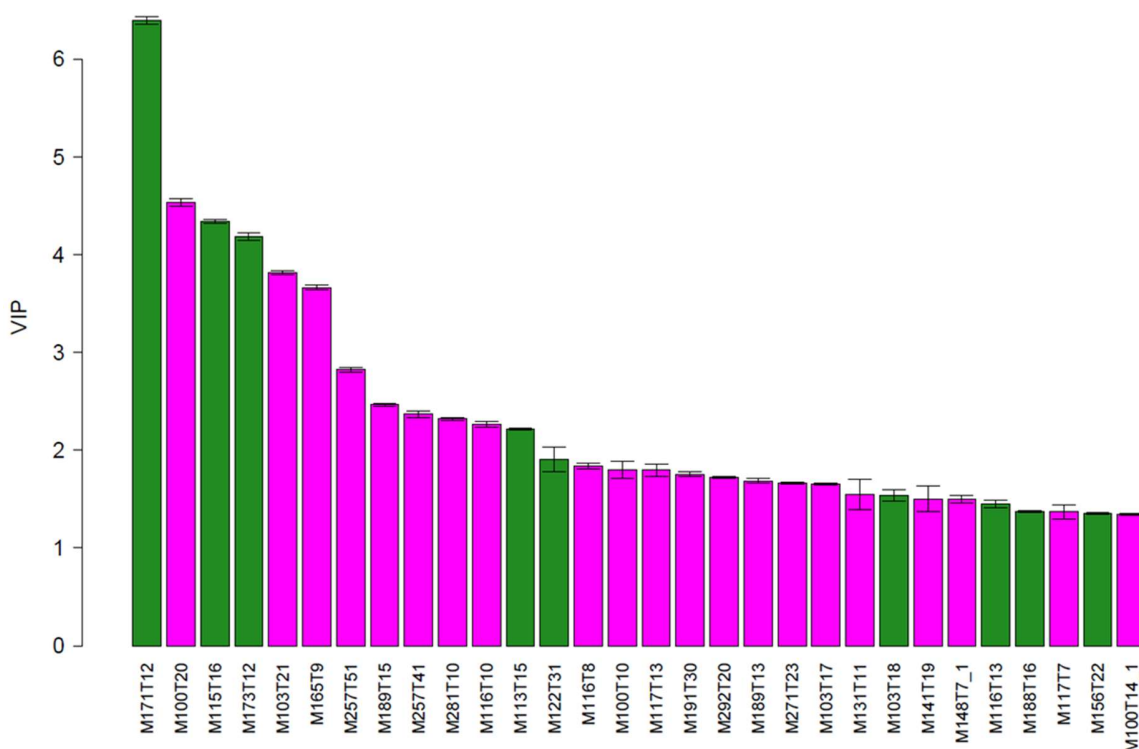
Obr. 20: VIP graf PLS-DA pro porovnání necílené GC-MS metabolomiky FUT a IUT. Metabolity zvýšené u FUT modře, u IUT růžově.



Obr. 21: VIP graf PLS-DA pro porovnání necílené GC-MS metabolomiky NT a IUT. Metabolity zvýšené u NT oranžově, u IUT růžově.

V analýze ureasových vzorků jsou přítomny i statisticky významné (VIP >2) artefakty necukerné povahy (Obr. 14, 15): dva v FUT (M100T10, M148T10\_2), jeden u IUT (M116T10). Nebylo možné je identifikovat pomocí NIST ani GOLM, ale na základě přítomnosti fragmentu s  $m/z$  100 se lze domnívat, že se jedná o látky obsahující aminoskupinu. Tyto artefakty mohou být ureasová residua, která se do analýzy dostávají při deproteinaci nebo degradaci ureasy. Na artefaktu M116T10 lze demonstrovat specifika metabolického přístupu. Z Obr. 15 je patrné, že se jedná o látku nízké intenzity. Ale protože v analýze FUT zcela chybí, je hodnocena jako statisticky významná. Látky s aminoskupinou by na LC-MS s ionizací pomocí elektrospreje (ESI) poskytovaly lepší odezvu a bylo by je možné lépe porovnat. Tyto artefakty však nejsou zahrnuty v cílené LC-MS metodě, tak jejich vliv na metabolom nelze pomocí LC-MS/MS hodnotit.

Na základě výsledků se lze domnívat, že přítomnost magnetického celulosového nosiče nezvyšuje množství artefaktů. Naopak, díky několikanásobnému promývání při přípravě imobilizované ureasy je množství kontaminantů při použití IUT nižší.



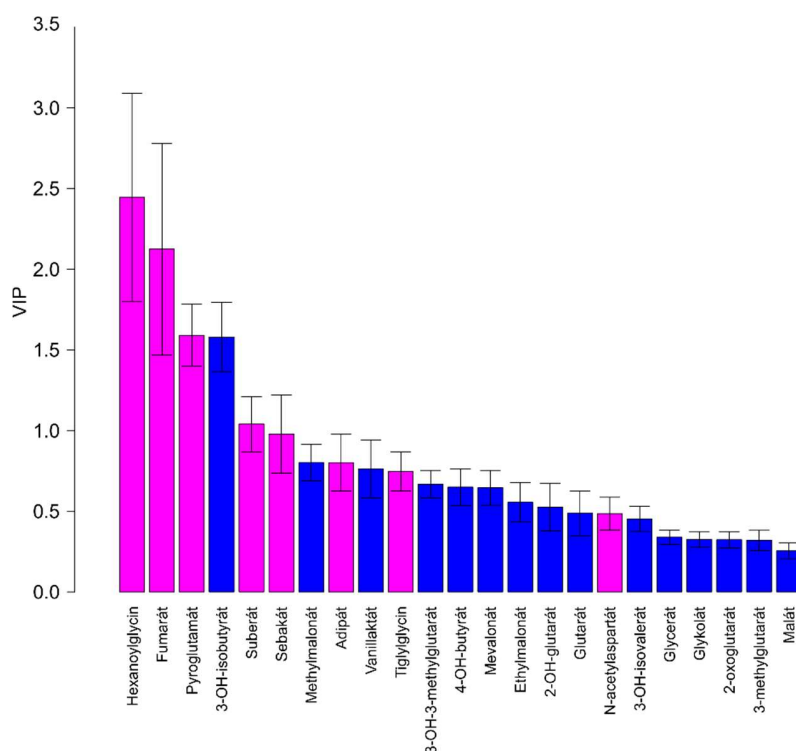
Obr. 22: VIP graf PLS-DA pro porovnání necílené GC-MS metabolomiky EX a IUT. Metabolity zvýšené u EX zeleně, u IUT růžově.

#### 8.5.4 Kvantitativní změny metabolomu

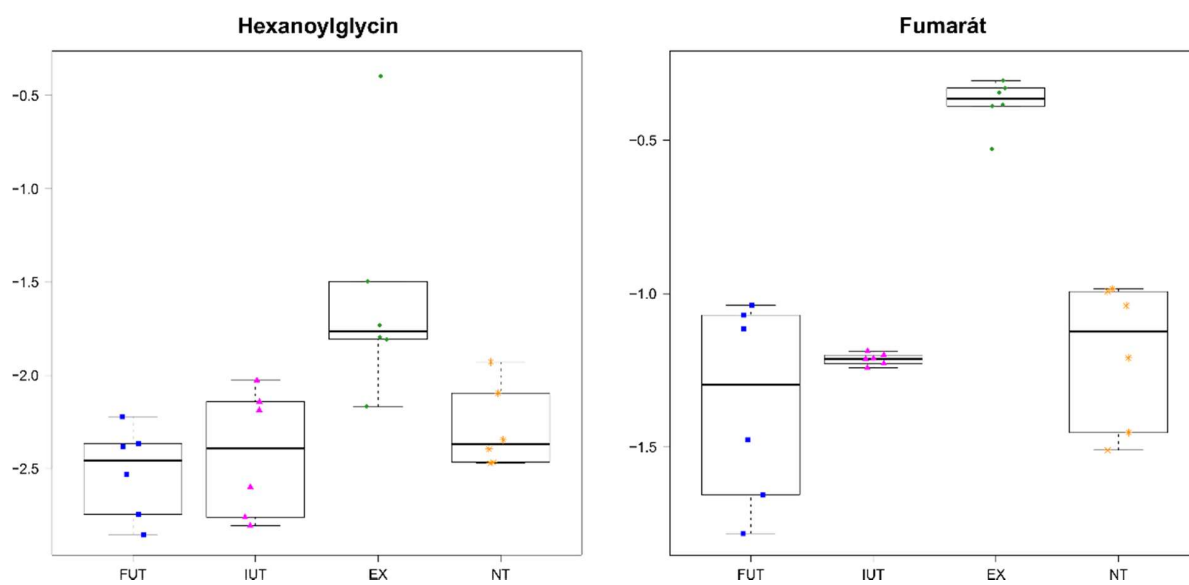
O ureasovém protokolu se obecně tvrdí, že vyvolává změny hladin některých metabolitů ve vzorcích [98]. Tyto změny metabolomu lze hodnotit pomocí cílené metabolomiky. Vliv rozdílných metod přípravy vzorku na intenzity konkrétních metabolitů lze demonstrovat pomocí poměru mediánů dvou skupin vzorků (jedná se o tzv. míru velikosti účinku, angl. „fold change“) a porovnáním analytické variability (CV) metabolitů.

Na základě PCA z cílené GC-MS metabolomiky lze usuzovat, že FUT a IUT mají stejný vliv na kvantitu metabolitů, zatímco každý z obou zbývajících (NT a EX) ovlivňuje intenzity jiným způsobem (viz kapitola 8.5.2). Detailní porovnání se provádí pomocí PLS-DA.

VIP graf cílené GC-MS metabolomiky FUT a IUT vzorků ukazuje dva metabolity s VIP>2 (Obr. 23), oba však vykazují vysokou variabilitu. Chybové úsečky zasahují pod VIP 2, a proto nejsou považovány za signifikantní. I z krabicových grafů je patrné, že se tyto analyty neliší, protože se jejich interkvartilová rozpětí překrývají (Obr. 24). Poměry mediánů se pohybují v rozmezí od 0,92 do 1,14 (Tab. 2). Cílená LC-MS (Obr. 25) odhalila pouze jediný rozdíl, a tím je zvýšený obsah „laktosy“ (metabolit C023), jak již bylo diskutováno v kapitole 8.5.3.



Obr. 23: VIP graf PLS-DA pro porovnání cílené GC-MS metabolomiky FUT a IUT. Metabolity zvýšené u FUT modře, u IUT růžově.

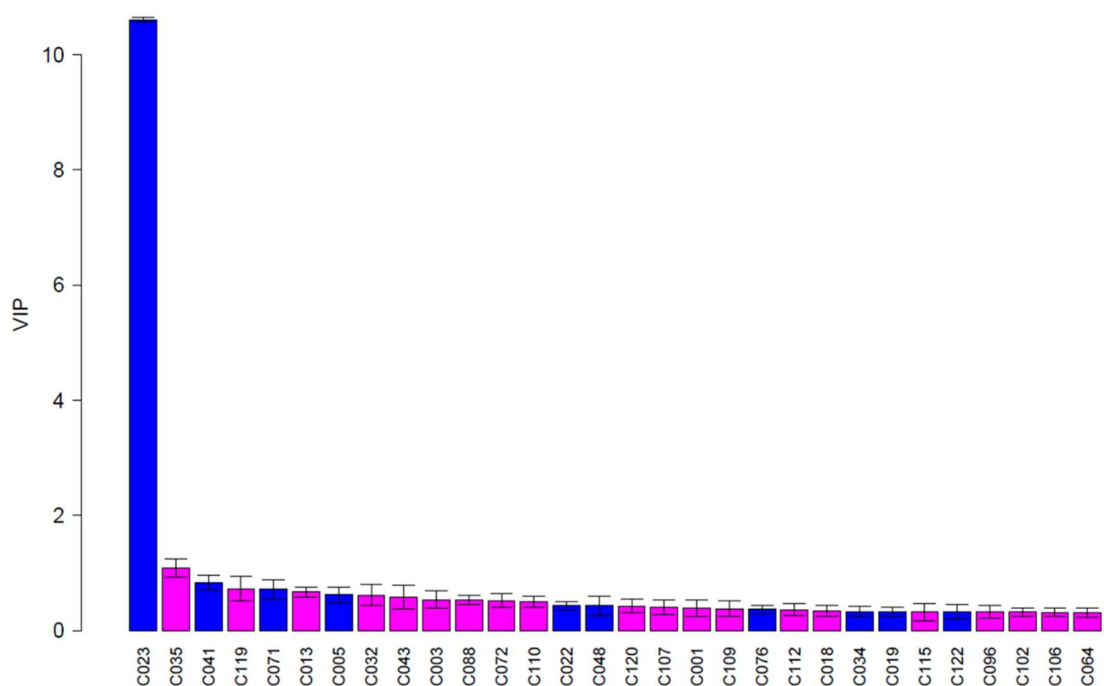


Obr. 24: Krabicové grafy zobrazující distribuci dat vzorků měřených pomocí cílené GC-MS metabolomiky. Zobrazeny jsou dva metabolity, které měly při porovnání FUT a IUT metodou PLS-DA  $VIP > 2$ .



Tab. 2: Retenční indexy (pro IUT) a poměry mediánů intenzit metabolitů cílené GC-MS analýzy. Tučně jsou vyznačeny metabolity s poměrem mediánů pod 0,6.

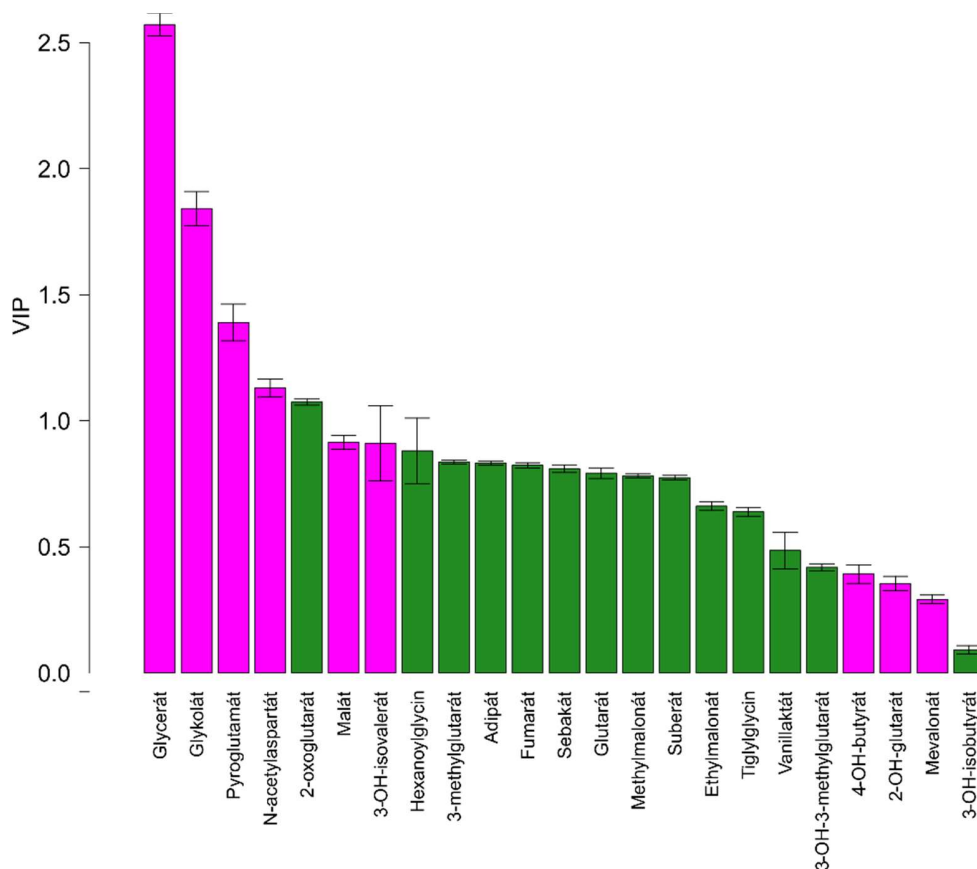
Název metabolitu	Retenční index	FUT/IUT	NT/IUT	EX/IUT
Glykolát	1070	1,09	0,67	<b>0,07</b>
3-hydroxyisobutyřát	1158	1,18	0,73	<b>0,47</b>
3-hydroxyisovalerát	1207	1,08	0,98	<b>0,23</b>
Methylmalonát	1213	1,14	0,90	1,00
4-hydroxybutyřát	1236	1,10	0,80	<b>0,32</b>
Ethylmalonát	1278	1,06	<b>0,23</b>	0,86
Glycerát	1329	1,08	0,92	<b>0,04</b>
Fumarát	1351	0,98	0,86	1,01
Mevalonát/mevalonolakton	1557/1362	1,12	<b>0,48</b>	<b>0,36</b>
Glutarát	1405	1,05	0,85	1,02
3-Methylglutarát	1425	1,05	0,81	1,03
Malát	1488	1,08	0,80	<b>0,19</b>
Adipát	1506	1,03	0,91	1,06
Pyroglutamát	1515	0,92	1,08	<b>0,11</b>
Tiglylglycin	1568	1,00	0,73	0,81
2-hydroxyglutarát	1576	1,07	0,77	<b>0,33</b>
3-hydroxy-3-methylglutarát	1600	1,11	0,79	0,73
2-oxoglutarát	1621	1,04	1,05	1,31
Hexanoylglycin	1635	0,93	0,77	0,92
N-acetylaspartát	1655	1,03	0,74	<b>0,15</b>
Suberát	1697	0,97	0,77	0,97
Sebakát	1893	1,01	0,80	1,01
Vanillaktát	2010	1,08	0,78	0,67
Průměr		1,05	0,79	0,36



Obr. 25: VIP graf PLS-DA pro porovnání cílené GC-MS metabolomiky FUT a IUT. Metabolity zvýšené u FUT modře, u IUT růžově. Metabolit C023 je laktosa. Seznam všech analytů je uveden v Tab. S1.

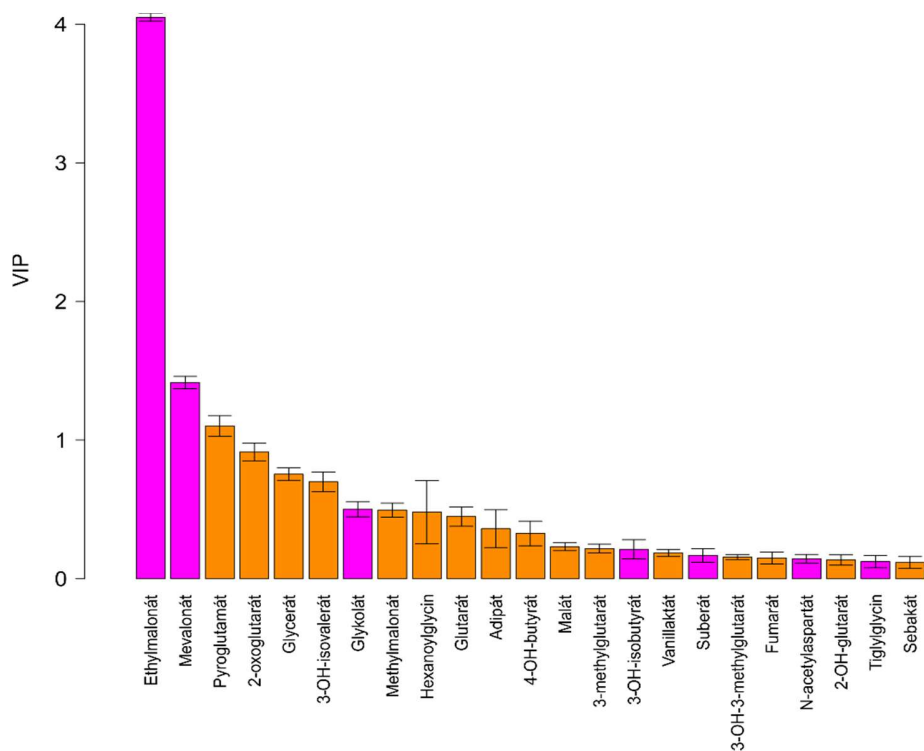
Při porovnání EX a IUT je nutné vzít v úvahu, že EX nebyl původně vyvinut pro odstranění močoviny, ale pro izolaci organických kyselin a glycinových konjugátů ze vzorku [108]. Výsledné složení vzorku proto záleží na zvoleném organickém rozpouštědle. Pokud je jím methanol, pak vzorek obsahuje hlavně lipidy [109]. Je-li jím voda, tak ve vzorku převažují cukry a hydroxysloučeniny. Při použití vody je přítomen také malý podíl středně polárních látek, zatímco lipidy zcela chybí [73]. Nejčastěji užívaným postupem je kyselá extrakce do ethylacetátu, proto byla využita i v této práci [20]. Metoda vyniká dobrou výtěžností (angl. „recovery“). Pro většinu diagnosticky významných organických kyselin a glycinových konjugátů se pohybuje v rozsahu 60-107 %, citrát, methylcitrát, N-acetylaspartát, 2-hydroxybutyrát a 3-hydroxybutyrát jsou však zatíženy nízkou extrakční účinností [20,110]. Jak můžeme vidět z poměru mediánů EX a IUT (Tab. 2), extrakční účinnost je u 11 z 23 metabolitů mimo rozmezí uvedené Tanakou a kol. (výtěžnost 60-107 % [20]). Až na jedinou výjimku, všechny dikarboxylové kyseliny a glycinové konjugáty toto kritérium splňují. Tou výjimkou je 2-oxoglutarát, který překročil uvedený limit. Patří mezi ketokyseliny, o nichž bylo doloženo, že je jejich odezva ve vzorcích naředěných na koncentraci kreatininu nižší než 2 mmol.l<sup>-1</sup> proměnlivá. Na základě VIP grafu je jediným

signifikantně odlišným metabolitem glycerát (Obr. 26). Jak již bylo zmíněno v kapitole 8.3.2, vzorky EX nelze připravit pro LC-MS.

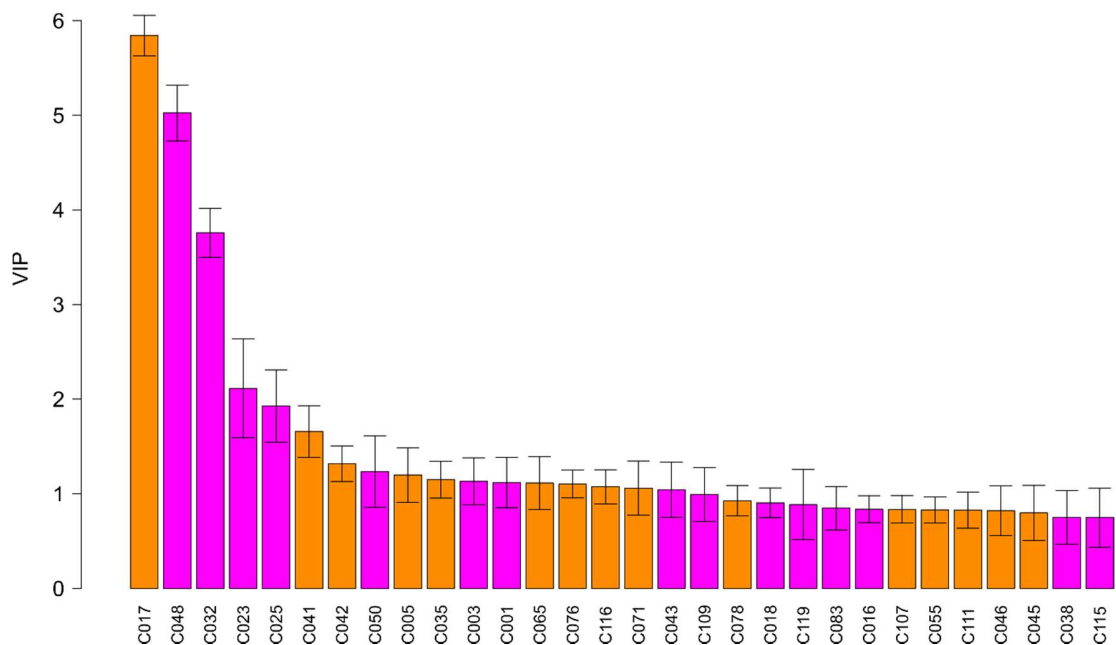


Obr. 26: VIP graf PLS-DA pro porovnání cílené GC-MS metabolomiky EX a IUT. Metabolity zvýšené v IUT růžově, u EX zeleně.

Analýza NT vzorků byla zatížena přítomností obrovského píku močoviny (Obr. 14). Ethylmalonát, který je na koloně zadržován stejně jako močovina, je nejvíce diskriminujícím metabolitem (Obr. 27). PCA LC-MS analýzy (Obr. 17) ukazuje, že vzorky ošetřené IUT poskytují podobné výsledky jako vzorky bez úpravy (NT). Dle VIP grafu (Obr. 28) však analýzu NT a IUT odlišují 3 metabolity: uracil (C017), cytidin (C048) a glukuronát (C032). Jelikož VIP graf porovnávací FUT a IUT indikuje pouze jediný signifikantně odlišný metabolit, je zřejmé, že tento jev je společný pro FUT a IUT. Jeho příčinu se však nepodařilo objasnit. Experimentálně byla vyloučena iontová suprese, vliv ureasou vyvolaného vysokého pH na rozpustnost či ionizaci (jako důsledek laktam-laktim tautomerie).



Obr. 27: VIP graf PLS-DA pro porovnání cílené GC-MS metabolomiky NT a IUT. Metabolity zvýšené u IUT růžově, u NT oranžově.



Obr. 28: VIP graf PLS-DA pro porovnání cílené LC-MS metabolomiky NT a IUT. Metabolity zvýšené u IUT růžově, u NT oranžově; C017 - uracil, C048 – cytidin, C032 – glukuronát.

### 8.5.5 Analytická variabilita

U GC-MS analýzy se testované metody přípravy vzorku liší ve variabilitě odezvy 23 cílových metabolitů v COA: pro FUT činí medián hodnot 5,0 %, pro IUT 2,5 %, pro EX 19,7 % a pro NT 11,4 % (Tab. 3). Hodnota CV pro FUT je srovnatelná s dříve publikovanými výsledky (6,2 % pro 250 metabolitů [63]). Nejvíce variabilní odezvu poskytuje hexanoylglycin, a to pro všechny testované přístupy (22-119 %). Příčinou této variability však není příprava vzorku, ale chromatografie. Koncentrace hexanoylglycinu ve vzorku je po naředění na koncentraci kreatininu  $1 \text{ mmol.l}^{-1}$  pouze  $9 \text{ }\mu\text{mol.l}^{-1}$ , což se blíží limitu detekce. Navíc hexanoylglycin (stejně jako ostatní glycinu) poskytuje chvostující píky. Nejvyšší variabilitu vykazuje EX jako výsledek rozdílné extrakční účinnosti. Tento fenomén je patrný i z Tab. 1, která uvádí výrobcem deklarované laboratorní rozmezí pro 23 metabolitů obsažených v COA. Hodnoty byly stanoveny rutinními laboratořemi, ve kterých se dosud používá převážně tradiční EX. Např. pro 3-hydroxy-3-methylglutarát je uveden rozsah hodnot od 5 do  $281 \text{ }\mu\text{mol.l}^{-1}$ . Nejnižší variabilita je nalezena pro IUT, pravděpodobně jako výsledek jednodušší přípravy vzorku. Jelikož v této metodě není deproteinační krok, vzorky se snáze lyofilizují. Přítomnost methanolu ve vzorcích připravených pomocí FUT způsobuje, že vzorky v lyofilizátoru prskají.

LC-MS analýza neukazuje na výrazné změny v abundancích metabolitů mezi testovanými přístupy (Tab. S2). Medián hodnot CV se pohybuje od 9,5 do 11,7 %. Nejvyšší variabilitu vykazuje laktosa (66,3 %), protože jako příměs komerční ureasy se v NT vzorcích vyskytuje pouze ve stopových koncentracích. Ve vzorcích připravených pomocí IUT se objevují zvýšená CV pro amino- a thiosloučeniny (např. methionin, 2-aminoadipát, pantothenát, karnitiny), což by mohlo poukazovat na možnou interakci těchto látek s povrchem magnetického celulosového nosiče. Tento jev však u GC-MS analýzy nebyl prokázán (u pyroglutamátu, tiglylglycinu ani N-acetylaspartátu). Může to však být způsobeno rozdílnou citlivostí obou technik.

Tab. 3: Porovnání reprodukovatelnosti odezvy (plocha píku) analytů pomocí cílené GC-MS v COA mezi metodami

Metabolit	IUT	FUT	NT	EX
Glykolát	2,42	4,85	12,61	37,30
3-hydroxyisobutyřát	3,46	3,30	13,50	20,07
3-hydroxyisovalerát	2,18	4,90	8,17	48,68
Methylmalonát	2,51	3,84	11,78	15,21
4-hydroxybutyřát	6,22	7,88	10,22	19,09
Ethylmalonát	6,16	4,98	11,57	17,90
Glycerát	1,95	4,40	15,83	23,75
Fumarát	1,79	26,48	26,16	18,07
Mevalonát	1,67	4,96	13,54	16,73
Glutarát	7,73	5,87	10,59	17,66
3-methylglutarát	2,17	4,06	8,95	17,49
Malát	1,51	3,16	10,20	22,80
Adipát	2,04	7,37	23,02	16,80
Pyroglutamát	1,98	6,40	15,25	46,35
Tiglylglycin	3,06	4,96	7,14	18,63
2-hydroxyglutarát	5,37	3,35	11,42	25,55
3-hydroxy-3-methylglutarát	1,50	4,37	8,01	16,49
2-oxoglutarát	1,54	5,35	6,82	19,80
Hexanoylglycin	29,85	21,87	26,72	119,06
N-acetylaspartát	3,08	5,48	6,30	28,09
Suberát	3,20	4,40	12,91	19,55
Sebakát	5,04	6,04	7,52	19,69
Vanillaktát	2,95	8,07	9,65	57,54
Medián	2,51	4,96	11,42	19,69

### 8.5.6 Reprodukovatelnost retenčních časů

Reprodukovatelnost retenčních časů ( $t_r$ ) je základní požadavek pro správné zarovnání píků u necílené metabolické analýzy a také pro identifikaci metabolitů v metabolických databázích (např. GOLM) pomocí retenčních indexů [111]. V tomto experimentu se reprodukovatelnost  $t_r$  stanovuje cíleným metabolickým přístupem, pro 23 metabolitů obsažených v COA u GC-MS, pro 123 metabolitů identifikovaných v témže materiálu pomocí LC-MS.

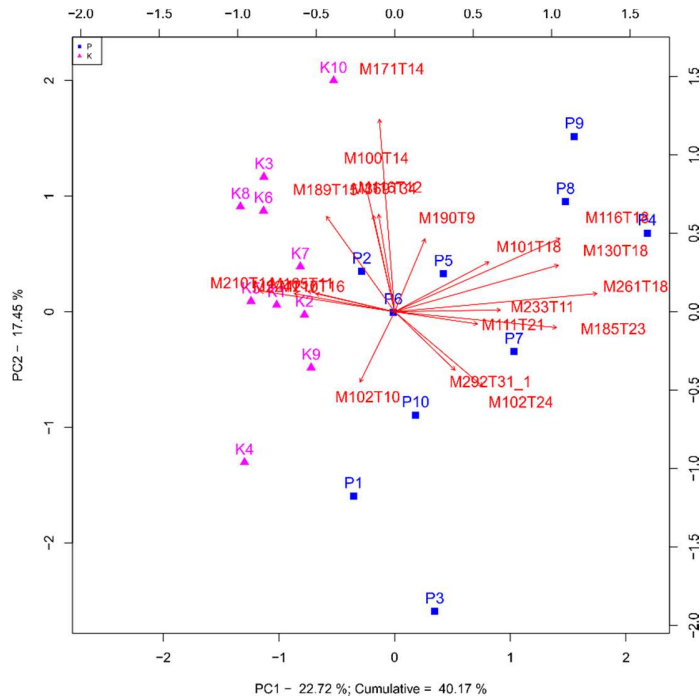
Retenční časy i retenční indexy jsou pro obě techniky i všechny metody stálé. Průměrná relativní směrodatná odchylka (CV) retenčních indexů (u GC-MS) byla nižší než 0,01 %. Největší rozdíly byly zaznamenány u NT vzorků, a to pro metabolity v retenčním okně močoviny (např. 0,15 % pro methylmalonát a pro 4-hydroxybutyrát), jako důsledek přesycení kolony. Variabilita retenčních časů u LC-MS je ve srovnání s GC-MS vyšší (průměrně 0,7 % pro každý z přístupů), ale pro všechny metody stejná. Obecně však lze říci, že je vyšší pro méně zadržované látky (2,4 % pro uracil s  $t_r$  2,4 min vs. 0,1 % pro ethylmalonát s  $t_r$  7,14 min).

### 8.5.7 Metabolický experiment s močemi pacientů s glutarovou acidurií I

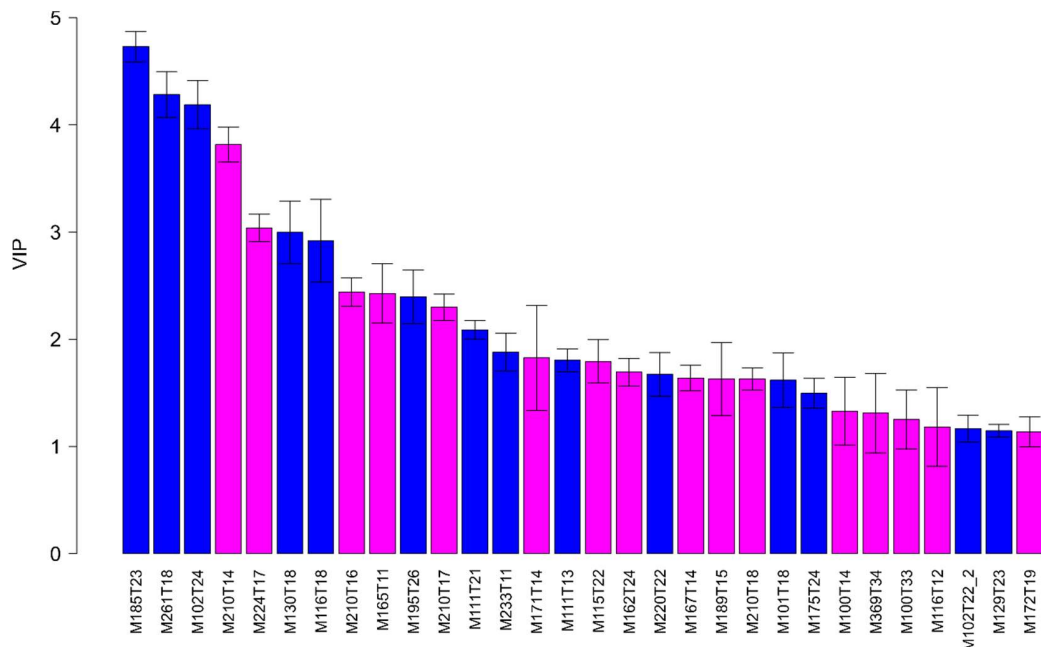
GAI je autozomálně-recesivní onemocnění způsobené deficitem glutaryl-CoA-dehydrogenasy (EC 1.3.99.7), díky kterému nejsou postižení jedinci schopni odbourávat lysin, hydroxylysin a tryptofan. Většinou se projevuje mozkovou atrofií způsobující změny v motorice a hybnosti, které jsou při zátěži organismu výraznější, často také i mentální retardaci [112]. V moči jsou přítomny zvýšené hladiny 3-hydroxyglutarátu a glutarátu [113].

Pro ověření aplikovatelnosti IUT pro metaboliku je imobilizovaná ureasa použita k přípravě moči 10 pacientů s GAI a 10 zdravých kontrol. Vzorky jsou analyzovány pomocí GC-MS, data jsou zpracována necíleným metabolickým přístupem. Ve vzorcích je nalezeno celkem 221 píků metabolitů. Pomocí PCA (Obr. 29) se vzorky dělí na 2 skupiny. V první skupině jsou zastoupeny pouze zdravé kontroly, ve druhé pacienti s GAI. U pacientů jsou v porovnání s kontrolami přítomny signifikantně zvýšené hladiny (Obr. 30) 3-hydroxyglutarátu (M185T23, RI 1575, shoda s knihovnou 968), tartarátu (M102T24, RI 1636, shoda s knihovnou 917), glutarátu (M261T18, RI 1403, shoda s knihovnou 880) a neznámé látky M111T13 (RI 1213). Snížené hladiny jsou pozorovány pro M210T14 (RI 1240), M224T17 (RI 1362), M210T16 (RI 1317), M210T17 (RI 1350) a M111T21 (RI 1507). Glutarát a 3-hydroxyglutarát jsou známými biomarkery onemocnění, tartarát je součástí léků podávaných pacientům s GAI. Neznámé látky M210T14, M224T17, M210T16, M210T17 a M111T21 jsou

komponenty moči zdravých jedinců, jež se nepodařilo identifikovat pomocí NIST ani GOLM databáze. Jejich identifikace bude předmětem další práce.



Obr. 29: Dvojný graf PCA nečíslené GC-MS analýzy moči pacientů s glutarovou acidurií I (P1-P10) a zdravých kontrol (K1-K10), připravených pomocí imobilizované ureasy.



Obr. 30: VIP graf PLS-DA nečíslené GC-MS analýzy moči pacientů s glutarovou acidurií I a zdravých kontrol. Metabolity zvýšené u GAI modře, u kontrol růžově.



## 8.6 Závěr

Ureasa imobilizovaná na magnetických částicích je novou technikou přípravy vzorku moči pro GC-MS analýzu. Ve srovnání s ostatními přístupnými přípravami (kyselá extrakce do ethylacetátu, volná ureasa, vzorek bez úpravy) vyniká snadností přípravy i reprodukovatelností výsledků (pro 23 organických kyselin a acylglycinů obsažených v referenční moči činí průměrné CV 2,5 %).

Chromatogramy vzorků připravených pomocí imobilizované ureasy jsou podobné přístupu za použití volné ureasy, v analýze je však méně artefaktů (jen 3 z původních 10). Je tomu tak nejspíš proto, že imobilizovaný enzym lze před aplikací vzorku snadno a opakovaně promýt. Vzorky připravené extrakcí poskytují zcela odlišné chromatografické záznamy. Ve srovnání s oběma ureasovými protokoly obsahují zbytky močoviny a také odlišné hladiny metabolitů (především kreatininu a glycerátu). Vzorky bez úpravy jsou u LC-MS srovnatelné s imobilizovanou ureasou, protože močovina se eluuje v mrtvém čase, a tak neovlivňuje cílové analyty. Při aplikaci GC-MS je chromatografický systém přetížen přítomností vysoké koncentrace močovinou, která vede ke změnám retence i odezvy metabolitů, zejména ethylmalonátu.

Pomocí metabolomického experimentu se vzorky pacientů s glutarovou acidurií I je ověřeno, že imobilizovanou ureasu lze použít i pro necílenou GC-MS metabolomiku. Chromatografická data je možné zpracovávat běžným způsobem. Hlavními diskriminujícími metabolity jsou biomarkery GAI, a tak lze usuzovat, že aplikace imobilizované ureasy neovlivňuje metabolom moči.

## 9 WOSEL jako nový způsob škálování 2D separačního prostoru

Dvojměrné chromatografické systémy jsou tvořeny dvěma separačními dimenzemi, které se liší ve svém rozsahu [43,114,115]. Např. v GC×GC může separace v první dimenzi trvat 3000 s, zatímco v druhé pouze 4 s, a tak se separace v dosažená v druhé dimenzi může jevit zanedbatelnou. Pokud by to byla pravda, pak by ale připojení další separační dimenze bylo zbytečnou komplikací a ne přínosem. Proto, aby se kladl stejný důraz na separaci dosaženou v první a druhé dimenzi, se retenční data před vyhodnocením efektivity dané separace škálují. Škálování má ale ještě další dva důsledky: tvoří normalizovaný retenční prostor (a tak je někdy označováno též jako normalizace) a odstraňuje nevyužitý či nevyužitelný separační prostor.

### 9.1 Způsoby škálování separačního prostoru

V pracích zabývajících se optimalizací dvojdímenzionálních systémů pomocí ortogonalit se škálování 2D separačního prostoru provádí dvěma různými způsoby, které lze podle hranic, které retenční prostor vymezují, označit jako void-max [43] a min-max [38]. Oba způsoby škálování jsou pro první dimenzi znázorněny na Obr. 31. U staršího z nich, void-max, je normalizovaný separační prostor je pro vymezen mrtvým časem dané dimenze a retenčním časem nejvíce zadržované složky vzorku použitého pro optimalizaci. Protože určení mrtvého času vyžaduje dodatečnou analýzu či výpočet, navrhl Gilar a kol. [38] zjednodušené min-max škálování, ve kterém je mrtvý čas nahrazen retenčním časem nejméně zadržované látky dané dimenze. Optimalizace 2D systému se však tímto krokem stává zcela závislou na vzorku použitém k optimalizaci. Jiný vzorek (obsahující jiné analyty) vymezí normalizovaný retenční prostor (nesrovnatelně) jinak.

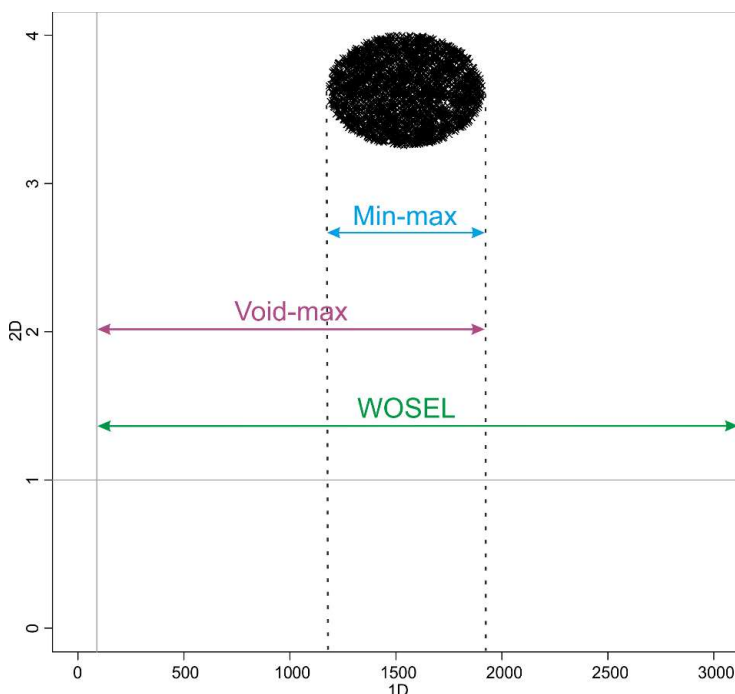
V případě cílené metabolomiky to nemusí představovat problém, protože je možné opakovaně namíchat stejnou směs standardů. V případě necílené metabolomiky, kdy všechny analyty přítomné ve vzorku nejsou předem známy, a navíc se přidává i vliv matrice, představuje takovéto škálování značnou komplikaci. Většinou se řeší tak, že se systém optimalizuje přímo pomocí daného biologického vzorku (např. směsná moč či plasma). Jedná se o neopakovatelný vzorek, jehož množství je omezeno. Ale ani tento vzorek neobsahuje analyty, které jsou důležité pro konkrétní experiment, protože biologické vzorky vykazují velkou interindividuální variabilitu v závislosti na aktuálním patofyziologickém stavu jedince, nutriční a dalších faktorech. Nejsprávnější by bylo použít k optimalizaci přímo QC vzorek daného experimentu. V případě vzácných onemocnění takovéto „plýtvání“ nepřipadá v úvahu. Optimalizace 2D systému je navíc časově náročná, většinou trvá několik dní. Vzorky pro daný experiment, už jednou

rozmražené kvůli přípravě QC vzorku, by se proto musely znovu zamrazit a po dobu optimalizace skladovat v mrazáku. Opakovaným rozmrazováním a delším skladováním se však metabolom vzorku mění [50]. A tak by i optimalizace provedená s konkrétním QC vzorkem nebyla se vzorkem 100% kompatibilní s experimentálními vzorky.

Cílem této části disertační práce je vyvinout škálování 2D prostoru, které by bylo nezávislé na konkrétním vzorku použitým k optimalizaci a co nejdříve odráželo původní separační prostor, a porovnat je pomocí syntetických (generovaných *in silico*) a reálných GC×GC chromatogramů se staršími škálovacími algoritmy. Nový přístup škáluje 2D separační prostor v souladu s analytickou metodou, tedy nezávisle na vzorku. Separační prostor je u něj vymezen mrtvým časem dané dimenze a posledním okamžikem sběru dat v dané dimenzi (u druhé dimenze je tento okamžik shodný s modulační periodou) (Rov. 11). K výpočtu ortogonality využívá celého separačního prostoru, proto se postup škálování nazývá WOSEL (z angl. „WhOle SEparation space scaLing“).

$$X(A) = \frac{t_r(A) - t_0}{t_r(\text{end}) - t_0} \quad (\text{Rov. 11}),$$

kde  $X(A)$  je normalizovaný retenční čas látky A,  $t_r(A)$  je retenční čas látky a v dané dimenzi,  $t_0$  je mrtvý čas dané dimenze,  $t_r(\text{end})$  je konec sběru dat v dané dimenzi.



Obr. 31: Různé způsoby škálování 2D separačního prostoru pro první dimenzi. Černý shluk představuje skupinu píků. Na osách je uveden retenční čas v sekundách.

## 9.2 Metody

### 9.2.1 Syntetické chromatogramy

Syntetické chromatogramy jsou v R-software (v. 3.3.0) vytvořeny tak, že tvoří oválný shluk obsahující 100 a 2500 datových bodů uvnitř separačního prostoru, který má v první dimenzi rozsah 89-3000 s (89 s je mrtvý čas první dimenze), ve druhé 1-4 s (1 s je mrtvý čas druhé dimenze). Data jsou generována do normalizovaného prostoru, který má v obou dimenzích rozsah 0 - 1, tak, že tvoří kruh o průměru 1,0 nebo 0,25 (Obr. 33, 34). Body uvnitř shluku mají rovnoměrné rozdělení, minimální vzdálenost bodů je 0,001. Souřadnice datových bodů jsou poté z normalizovaného retenčního prostoru přepočítány na reálný prostor (v obrázku označený jako „Original“). Protože se rozsah první a druhé dimenze liší, vzniká v původním separačním prostoru oválný shluk.

K ověření vlivu škálování na ortogonalitu systému se využívá 4 typů syntetických chromatogramů:

1. jednotkový shluk (průměr shluku je v normalizovaném retenčním prostoru roven 1,0)
2. čtvrtinový shluk (průměr shluku je v normalizovaném retenčním prostoru roven 0,25) umístěný v levém horním rohu
3. čtvrtinový shluk umístěný v pravém horním rohu
4. čtvrtinový shluk umístěný nahoře uprostřed

Aby se ověřila validita každého z modelů (prostřednictvím směrodatných odchylek), je každý typ chromatogramu generován pomocí daného algoritmu založeného na náhodném rozdělení 50×, a to jak pro 100 tak i pro 2500 bodů.

### 9.2.2 Reálné chromatogramy

K ověření výsledků syntetických chromatogramů je použito reálných GC×GC separací močí, plasmy a nafty. Mrtvé časy jsou určeny pomocí software Column calculator a Chroma-TOF (v. 4.51.6.0, LECO, St. Joseph, Michigan, USA), a to na základě rozměrů použitých kolon, průtoku helia a teplotního programu). Mrtvý čas u syntetických chromatogramů je odvozen z reálných separací metabolomických vzorků, připravených ethoximací a trimethylsilylací. Pro přípravu močí se využívá kyselá extrakce do ethylacetátu [116], u plasmy pak deproteinace pomocí methanolu, odfoukání pod dusíkem, ethoximace a trimethylsilylace [117]. Detailní postup přípravy vzorků i jejich chromatogramy jsou uvedeny v kapitole 11.

### 9.2.3 Výpočet ortogonalit

Poloha každého bodu (vrcholu píku) v 2D separačním prostoru je určena dvěma souřadnicemi: retenčním časem první a druhé dimenze. Těmto souřadnicím se říká retenční koordináty a jsou vstupními daty pro výpočet ortogonalit systému. Ortogonalita se v této práci počítá několika různými přístupy vybranými tak, aby z každé skupiny přístupů o stejném principu byl vybrán právě jeden: Pearsonova korelace (PC) [34], podmíněná entropie (CE) [35], konvexní obálka (CH) [36], modelování (MA) [42], geometrické pokrytí separačního povrchu (SC) [38], metoda nejbližšího souseda (NN) [40] a hvězdičkové rovnice (AE) [41]. Principy jednotlivých přístupů jsou uvedeny v úvodní části této disertační práce. Retenční koordináty píků se před výpočtem ortogonalit škálují třemi různými způsoby, uvedenými jako Rov. 9-11. Všechny výpočty se provádějí v R-software.

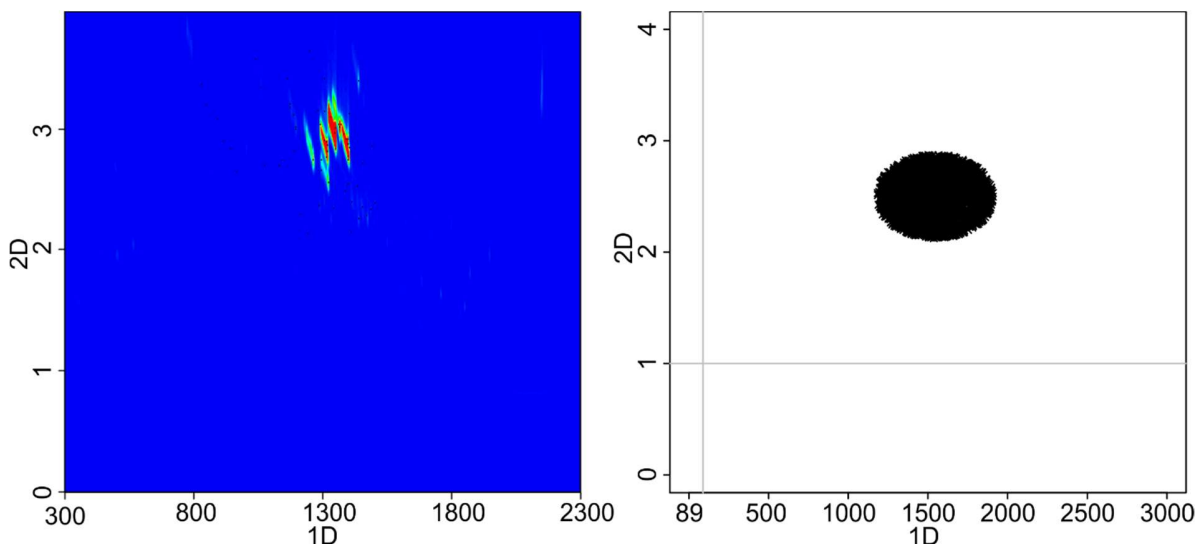
## 9.3 Výsledky a diskuse

Jednotlivé přístupy škálování 2D separačního prostoru se od sebe odlišují způsobem vymezení separačního prostoru (Rov. 9-11): začátek je definován pomocí  $t_0$  nebo  $t_r(min)$ , konec  $t_r(max)$  či  $t_r(end)$ . Zatímco  $t_r(min)$ ,  $t_r(max)$  lze snadno odečíst z každého chromatogramu a  $t_r(end)$  z parametrů metody, k určení  $t_0$  je nezbytné provést dodatečnou analýzu. U GC×GC se k stanovení mrtvého času polárních kolon používají alkany, pro nepolární kolony alkoholy [118]. Také u LC×LC volba nezadržované látky závisí na zvoleném separačním systému. Například pro reverzní fázi lze použít anorganickou sůl nebo jej určit z analýzy iontové suprese. Další možností přibližně přesného určení mrtvého času je požití výpočetního software (tzv. „void-time calculator“), který jej určí z fyzikálně-chemických vlastností analytického systému (např. teplotní program, parametry kolon, průtok mobilní fáze). Tato teoretická hodnota se od experimentální liší jen zanedbatelně.

### 9.3.1 Volba syntetických modelů

V dříve publikovaných pracích zabývajících se stanovením ortogonalit dvojdimenzionálních systémů se využívá celé řady syntetických modelů. Jejich 2D prostor protínají nejrůznější čtverce, obdélníky, trojúhelníky, linie či pásy [41,119,120], které ne zcela věrně napodobují chromatogramy reálných vzorků. Při tvorbě syntetických modelů pro účely této práce se výchozím bodem staly reálné GC×GC chromatogramy plasmy a moči, na kterých strukturně podobné látky tvoří oválné shluky (Obr. 32), jaké byly dříve popsány také Giddingsem v jeho původní práci pojednávající o konceptu dvojdimenzionálních separací (Obr. 7A) [30]. Zatímco v 1D separacích tvoří tyto látky široké píky, ve 2D separacích jsou díky vyšší frakcionační kapacitě systému patrné shluky ve tvaru elipsy s jemnou vnitřní

strukturou. Kromě obvyklých modelů obsahujících 100 datových bodů [38] je k testování použito i modelů obsahujících 2500 bodů, které jsou bližší v necíleným metabolickým experimentům. Tyto modely navíc vykazují velmi nízkou variabilitu výsledků [39,41].



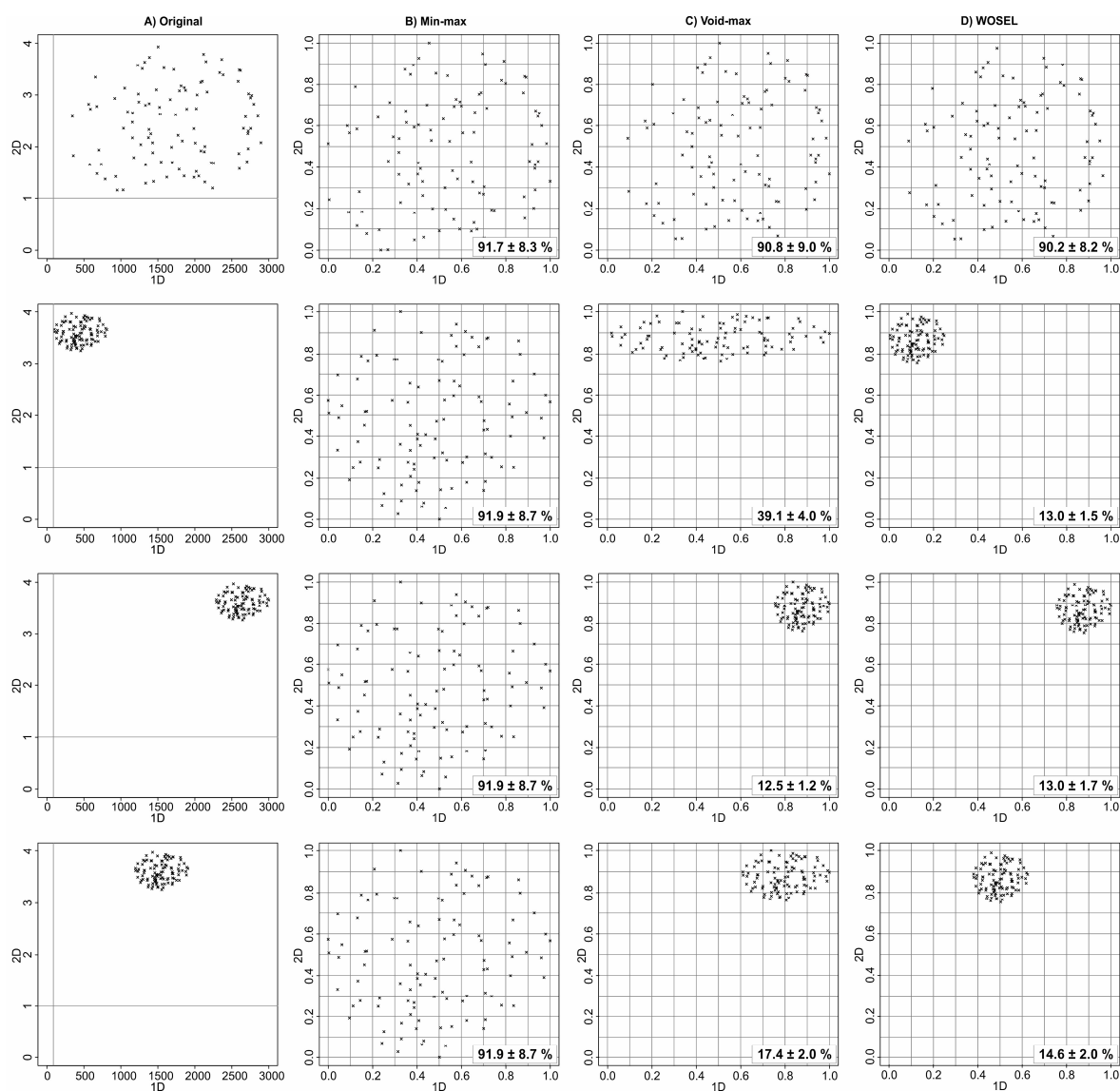
Obr. 32: Tvorba syntetických modelů (vpravo) na základě reálných separací lidské plasmy (vlevo). Na osách je uveden retenční čas v sekundách.

### 9.3.2 Vliv metody škálování 2D prostoru na výslednou ortogonalitu

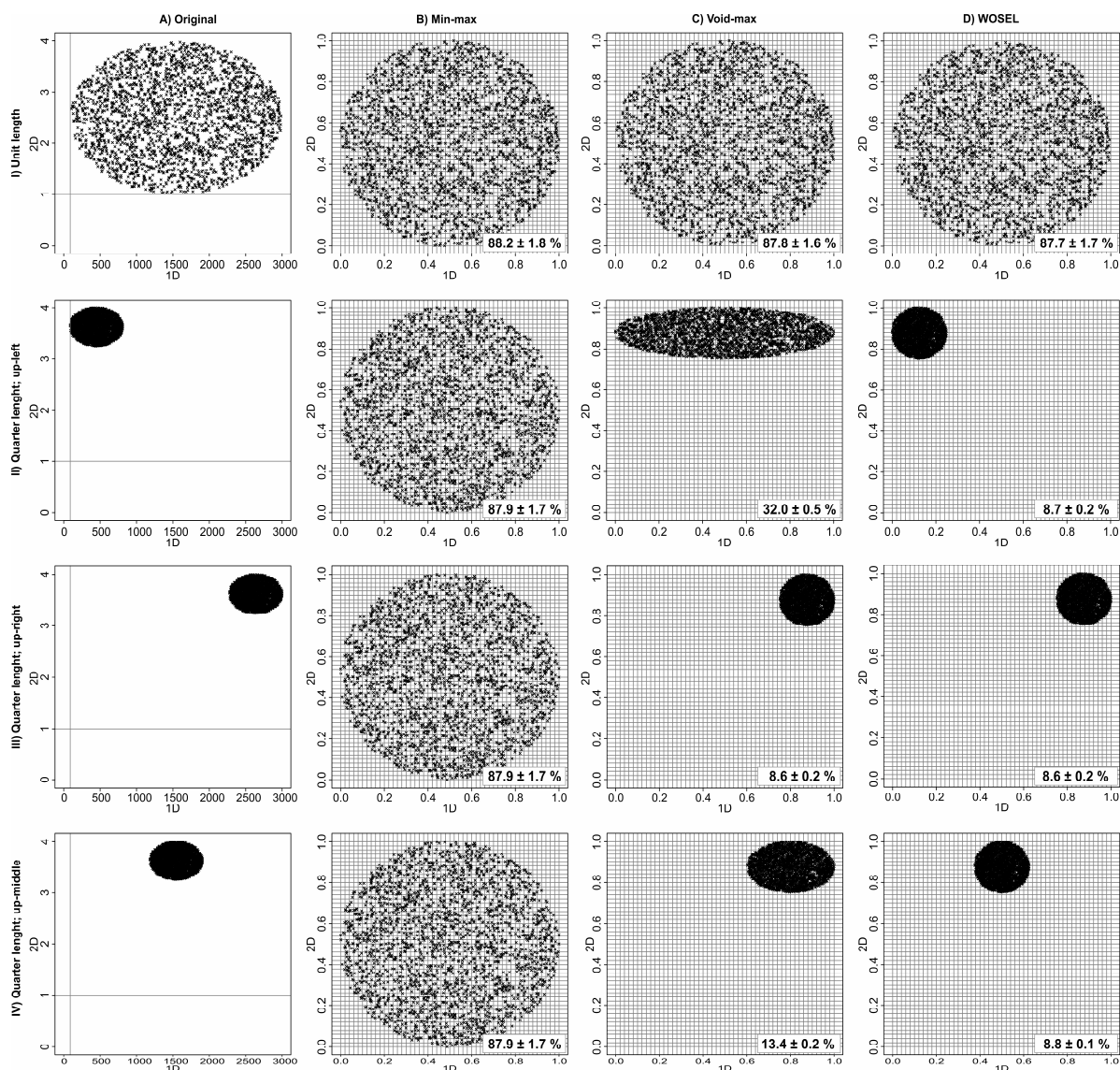
Při vyhodnocování optimálních dvojdimenzionálních separací dodnes hraje velkou roli vzhled chromatogramu (tzv. „percieved quality of chromatogram“) [121]. Jako dobrá separace bývá označována taková, ve které je využita velká část separačního prostoru a data tento prostor rovnoměrně pokrývají. Proto jsou výsledky prezentovány nejen v podobě číselných výsledků (na Obr. 33-36 jsou v pravém dolním rohu zobrazeny výsledky ortogonality pomocí SC), ale je zde uveden vzhled výchozího chromatogramu (označený jako „Original“) a také příslušných normalizovaných retenčních prostorů, které vznikají z výchozího aplikací jednotlivých přístupů škálování: min-max (Rov. 10), void-max (Rov. 9) a WOSEL (Rov. 11).

Z Obr. 33-35 vyplývá, že zvolený způsob škálování má zásadní vliv na výslednou ortogonalitu, zejména u vzorků, jejichž složky mají v 2D chromatogramu malou disperzi. Syntetické modely se 100 (Obr. 33) i 2500 body (Obr. 34) ukazují, že nejvýraznější rozdíly jsou patrné u vzorků, které obsahují skupinu látek vysoce zadržovaných v první i druhé dimenzi (shluk datových bodů je umístěn v pravém horním rohu chromatogramu). Vypočítaná ortogonalita nejvíce odpovídá vzhledu chromatogramu při

škálování pomocí WOSEL (který ukazuje, že je využito méně než 15 % plochy, Obr. 32). Starší způsoby vedou k nadhodnocování ortogonalit chromatogramů (min-max poskytl hodnotu vyšší než 91 %), protože do normalizovaného retenčního prostoru převádějí jen velmi malou část prostoru původního.



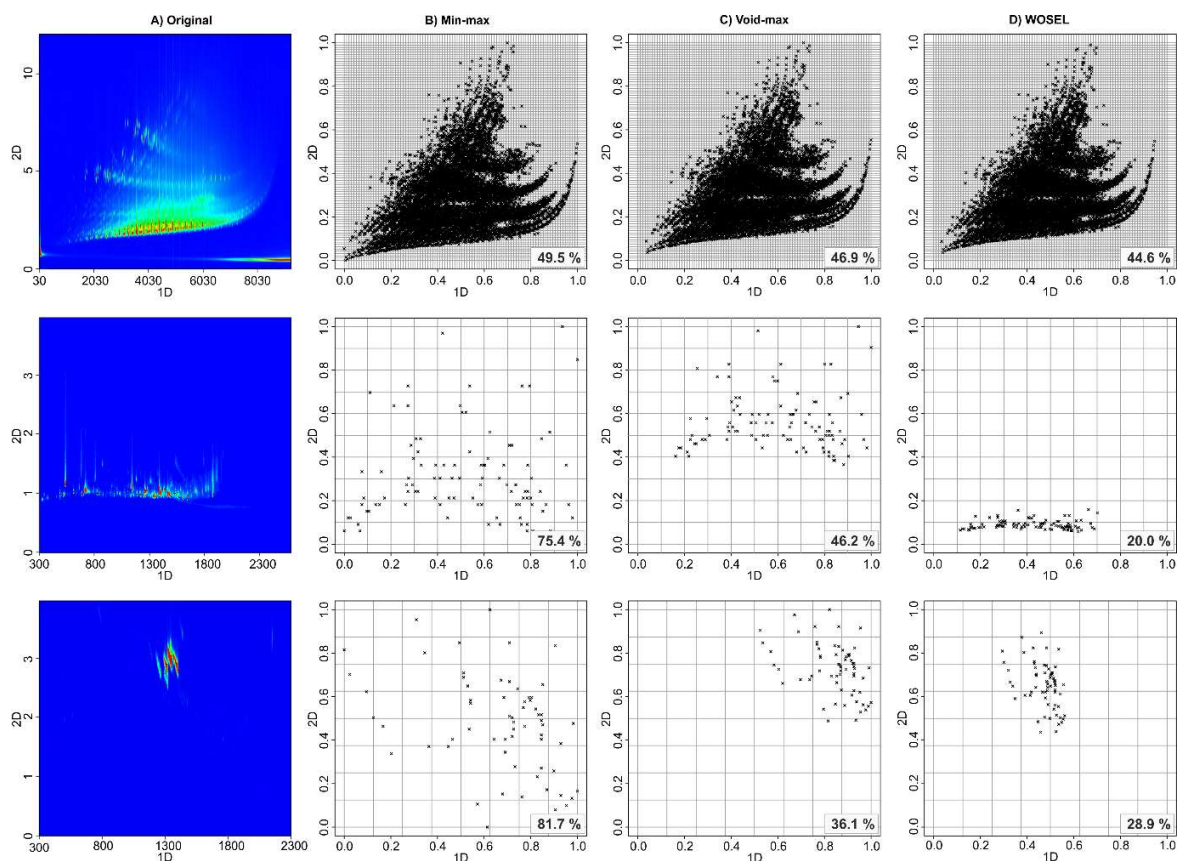
Obr. 33: Vliv metody škálování, demonstrováný pomocí modelů se 100 datovými body, na výslednou ortogonalitu, spočítanou pomocí SC (hodnota uvedená v pravém dolním rohu, společně se směrodatnou odchylkou pro 50 modelů generovaných stejným algoritmem).



Obr. 34: Vliv metody škálování, demonstrováný pomocí modelů se 2500 datovými body, na výslednou ortogonalitu, spočítanou pomocí SC (hodnota uvedená v pravém dolním rohu, společně se směrodatnou odchylkou pro 50 modelů generovaných stejným algoritmem).



Stejné závěry vyplývají i ze škálování reálných vzorků GC×GC separací nafty, moči a plasm (Obr. 35). Typ škálování má na ortogonalitu vzorku nafty jen velmi malý vliv, protože píky vyplňují velkou část separačního prostoru. Zatímco u separace moči a plasmy jsou patrné rozdíly mezi jednotlivými přístupy. Stejně jako v případě modelů (Obr. 33 a 34) hodnoty ortogonalit získané pomocí WOSEL nejvíce odpovídají vzhledu původního chromatogramu.



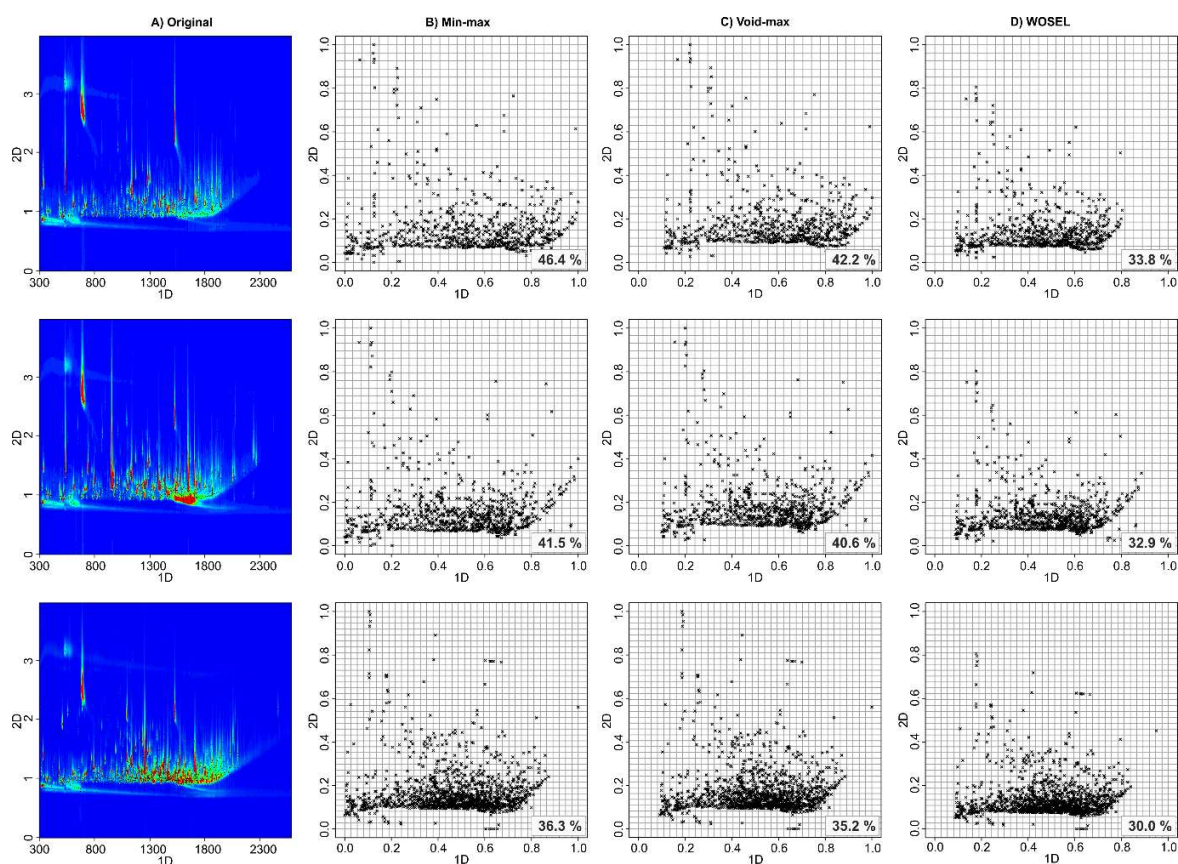
Obr. 35: Vliv metody škálování, demonstrováný pomocí GC×GC separace nafty (nahore), moči (uprostřed) a plasm (dole), na výslednou ortogonalitu, spočítanou pomocí SC (hodnota uvedená v pravém dolním rohu).

Při pohledu na Obr. 35 se nabízí otázka, proč je v separaci moči a plasmy ponechána tak velká část separačního prostoru prázdná. Vypadá to jako základní analytická chyba v optimalizaci. Ale není tomu tak. Při analýze metabolomických, klinických, forensních a toxikologických vzorků je nutné brát v úvahu fakt, že ne všechny látky, pro které byla separace optimalizována, musí být v konkrétním vzorku přítomny. V analýze však pro ně musí být ponechán prostor. Tento jev je demonstrován na příkladu mevalonové acidurie. Hlavními diagnostickými markery tohoto onemocnění je kyselina mevalonová a cholesterol. V GC×GC analýze se při tradičním kolonovém uspořádání (BPX5/BPX50) jedná o vysoce zadržované látky v první (cholesterol 1TMS) a druhé dimenzi (mevalonolakton 1TMS). Přestože se jedná o velmi vzácné onemocnění (dosud bylo doloženo přibližně 30 případů) [122], v analýze musí být ponechán prostor pro jejich zachycení.

### 9.3.3 Vliv vzorku použitého pro optimalizaci na škálování

Z rovnic 9-11 vyplývá, že normalizovaný retenční prostor, ve kterém se počítá ortogonalita, je, v případě min-max škálování zcela a v případě void-max částečně, vymezen analyty vzorku. Použije-li se k optimalizaci jiný vzorek stejného typu (moč jiného pacienta), je ortogonalita počítána v normalizovaném retenčním prostoru, do kterého byl převedena jinak velká část původního chromatogramu (Obr. 36).

Proto se výsledky ortogonality liší více, než je dáno samotnou přítomností či absencí analytů ve vzorku. Největší rozdíly mezi jednotlivými močemi jsou patrné při použití min-max škálování (přibližně 10 %), menší při void-max (7 %), nejmenší při WOSEL (méně než 4 %), přestože všechny tři vzorky mají srovnatelnou matici a liší se pouze v přítomnosti několika (diagnosticky významných) metabolitů. Tyto metabolity mají unikátní kombinaci funkčních skupin, a tak se v separačním prostoru vyskytují na novém místě. I když se normalizovaný retenční prostor nezmění, další bin je vyplněn píkem. Proto hodnota ortogonality nemůže být u všech močí ani za použití WOSEL zcela stejná, ale „jen“ srovnatelná, jako je tomu v případě uvedeném na Obr. 36.



Obr. 36: Vliv různého typu škálování na optimalizaci GCxGC systému pomocí 3 různých vzorků moče: nahore zdravá kontrola, uprostřed pacient s GAI, dole pacient v metabolickém rozvratu.

### 9.3.4 Vliv způsobu škálování na různé přístupy výpočtu ortogonalit

Z Obr. 33-36 vyplývá, že typ škálování ovlivňuje výsledky ortogonalit spočítané pomocí SC. Tento způsob výpočtu ortogonalit je pro demonstraci zvolen proto, že je všeobecně známý, často diskutovaný a běžně používaný [119–121,123,124]. Aby se prokázalo, zda typ škálování ovlivňuje i výsledky ortogonalit získané pomocí ostatních přístupů, byly retenční koordináty syntetických modelů použity k výpočtu ortogonalit pomocí jejich algoritmů. Výsledky shrnuté v Tab. 4 a Tab. S2 ukazují, že použitý typ škálování ovlivňuje všechny přístupy k výpočtu ortogonalit, s výjimkou PC, která si data v prvním kroku výpočtu sama škáluje (od každé hodnoty odečte průměr a pak ji vydělí směrodatnou odchylkou), a tak je nezávislá na použitém škálování [125]. Rozdíly ve výsledcích ortogonalit mezi jednotlivými způsoby škálování jsou překvapivě mnohem větší než mezi jednotlivými přístupy (blíže viz kapitola 10).

Tab. 4: Průměrné hodnoty ortogonalit jednotlivých typů modelů s 2500 body získané různými přístupy výpočtu ortogonalit  $\pm 2SD$  pro 50 modelů stejného typu. Hodnoty jsou uvedeny v %, s výjimkou AN a HN, jejichž výsledkem jsou průměrné nejmenší vzdálenosti mezi nejbližšími sousedy.

Typ modelu	Způsob výpočtu	Min-max	Void-max	WOSEL
	SC	88,2 $\pm$ 1,8	87,8 $\pm$ 1,6	87,7 $\pm$ 1,6
	MA	88,2 $\pm$ 1,8	87,8 $\pm$ 1,6	87,7 $\pm$ 1,6
	CE	85,0 $\pm$ 0,7	85,0 $\pm$ 0,6	85,1 $\pm$ 0,3
	PC	98,7 $\pm$ 1,9	98,7 $\pm$ 1,9	98,7 $\pm$ 1,9
	AE	79,4 $\pm$ 2,3	78,7 $\pm$ 2,2	78,3 $\pm$ 2,2
	CH	78,2 $\pm$ 0,7	77,6 $\pm$ 0,6	77,1 $\pm$ 0,2
	AN	0,0117 $\pm$ 0,0002	0,0116 $\pm$ 0,0002	0,0116 $\pm$ 0,0002
	HN	0,0078 $\pm$ 0,0006	0,0078 $\pm$ 0,0006	0,0078 $\pm$ 0,0006
	SC	87,9 $\pm$ 1,7	32,0 $\pm$ 0,5	8,7 $\pm$ 0,2
	MA	87,9 $\pm$ 1,7	32,0 $\pm$ 0,5	8,7 $\pm$ 0,2
	CE	85,0 $\pm$ 0,6	91,6 $\pm$ 0,7	95,3 $\pm$ 0,6
	PC	98,8 $\pm$ 1,9	98,8 $\pm$ 1,9	98,8 $\pm$ 1,9
	AE	79,1 $\pm$ 2,2	28,7 $\pm$ 1,0	4,9 $\pm$ 0,2
	CH	78,2 $\pm$ 0,7	19,4 $\pm$ 0,1	4,8 $\pm$ 0,0
	AN	0,0116 $\pm$ 0,0002	0,0058 $\pm$ 0,0001	0,0029 $\pm$ 0,0000
	HN	0,0078 $\pm$ 0,0006	0,0039 $\pm$ 0,0002	0,0019 $\pm$ 0,0002
	SC	87,9 $\pm$ 1,7	8,6 $\pm$ 0,2	8,6 $\pm$ 0,2
	MA	87,9 $\pm$ 1,7	8,6 $\pm$ 0,2	8,6 $\pm$ 0,2
	CE	85,0 $\pm$ 0,6	95,3 $\pm$ 0,6	95,3 $\pm$ 0,6
	PC	98,8 $\pm$ 1,9	98,8 $\pm$ 1,9	98,8 $\pm$ 1,9
	AE	79,1 $\pm$ 2,2	4,9 $\pm$ 0,2	4,9 $\pm$ 0,2
	CH	78,2 $\pm$ 0,7	4,8 $\pm$ 0,0	4,8 $\pm$ 0,0
	AN	0,0116 $\pm$ 0,0002	0,0029 $\pm$ 0,0000	0,0029 $\pm$ 0,0000
	HN	0,0078 $\pm$ 0,0006	0,0019 $\pm$ 0,0002	0,0019 $\pm$ 0,0002
	SC	87,9 $\pm$ 1,7	13,4 $\pm$ 0,2	8,8 $\pm$ 0,1
	MA	87,9 $\pm$ 1,7	13,4 $\pm$ 0,2	8,8 $\pm$ 0,1
	CE	85,0 $\pm$ 0,6	94,5 $\pm$ 0,6	95,2 $\pm$ 0,6
	PC	98,8 $\pm$ 1,9	98,8 $\pm$ 1,9	98,8 $\pm$ 1,9
	AE	79,06 $\pm$ 2,2	8,3 $\pm$ 0,3	4,9 $\pm$ 0,1
	CH	78,2 $\pm$ 0,7	7,7 $\pm$ 0,0	4,8 $\pm$ 0,0
	AN	0,0116 $\pm$ 0,0002	0,0037 $\pm$ 0,0000	0,0029 $\pm$ 0,0000
	HN	0,0078 $\pm$ 0,0006	0,0025 $\pm$ 0,0002	0,0019 $\pm$ 0,0002

### 9.3.5 Specifika vymezení separačního prostoru pomocí WOSEL

Vymezení separačního prostoru pomocí WOSEL je závislé na analytické metodě. Není proto nezbytné mít normalizovaný retenční prostor vymezen vždy jen pomocí  $t_0$  a  $t_r(end)$ . Začátek separačního prostoru může být v první dimenzi místo  $t_0$  určen např. začátkem sběru dat, pokud první minuty analýzy nejsou kvůli masivní eluci rozpouštědel zaznamenávány a detekce analytů by tam nebyla možná. U některých GC×GC separací se k analýze využívá také mrtvého času druhé dimenze, do kterého díky fenoménu anglicky nazývanému „wrap-around“ přecházejí látky vysoce zadržované v druhé dimenzi. V tomto případě je konec druhé dimenze dán  $t_0$ .

## 9.4 Závěr

WOSEL představuje nový způsob škálování 2D separačního prostoru, který do normalizovaného retenčního prostoru převádí celý prostor vymezený analytickou metodou (v obou dimenzích ohraničený mrtvým časem a koncem metody). Na rozdíl od starších metod škálování je díky tomu rozsah separačního prostoru nezávislý na vzorku použitém k optimalizaci.

Při škálování vzorků s velkým pokrytím separačního prostoru poskytuje WOSEL srovnatelné výsledky jako starší metody škálování. Pro vzorky s nižším pokrytím separačního prostoru však dává hodnoty odpovídající vzhledu chromatogramu. Umožňuje optimalizaci analytických systémů i pro analyty, které nejsou ve vzorku momentálně přítomny, přestože analytická metoda pro ně v separaci ponechává prostor.

## 10 Význam dimenzionality vzorku pro výpočet ortogonalit 2D separací

Dimenzionalita vzorku je souhrnem všech nezávislých vlastností vzorku, které mají vliv na jeho chromatografické chování. Většina vzorků jsou vzorky vysokodimenzionální, jejichž dimenzionalita je mnohonásobně vyšší než dimenzionalita použitého chromatografického systému. Píky těchto vzorků vyplňují bez jakéhokoli řádu celý 2D separační prostor. Oproti tomu vzorky nízkodimenzionální mají srovnatelnou dimenzionalitu jako analytický systém. Jejich píky vyplňují jen část chromatogramu. Formují se do skupin podle příslušnosti do chemických tříd. Poloha píku v rámci skupiny je dána pozicí funkční skupiny ve struktuře dané látky [45].

### 10.1 Lokální a globální ortogonalita

Ortogonalita je parametr, který se používá k optimalizaci chromatografických 2D systémů. Počítá se pomocí celé řady přístupů, pracujících na rozličných principech. Většinou nabývá hodnot od 0 do 100 %. Výjimku tvoří přístup založený na principu nejbližšího souseda (NN), jehož výsledkem je průměrná vzdálenost každého bodu chromatogramu a jeho nejbližšímu sousedovi v normalizovaném retenčním prostoru.

Obecně lze říci, že čím vyšší je ortogonalita (resp. průměrná vzdálenost NN), tím větší přínos má druhá dimenze pro celkovou separaci. Ortogonalita systému se hodnotí ze dvou různých hledisek, tzv. aspektů. Lokální ortogonalita popisuje rovnoměrnost pokrytí separačního prostoru, zatímco globální ortogonalita rozsah využitého separačního prostoru. V současné době neexistuje deskriptor, který by uměl zachytit obě hlediska ortogonalit. Proto některé deskriptory hodnotí pouze míru nahloučení dat, jiné zase zohledňují, kolik separačního prostoru je pokryto datovými body. Obě hlediska separace jsou však stejně důležitá, proto Schure a Davis v předchozí práci zabývající se porovnáním způsobů výpočtu ortogonalit zkombinovali nejlepší deskriptor globální (CH) a lokální ortogonalit (jednu z metod geometrického pokrytí separačního prostoru, založenou na počítání obsazených binů [126]) tak, že je mezi sebou vynásobili. Tento kombinovaný deskriptor dle závěru jejich práce [121] poskytuje výsledky, které nejlépe odpovídají vzhledu chromatogramů. Ve své práci však u všech přístupů, vyjma CH, použili min-max škálování, které zkresluje reálný separační prostor. CH tak v jejich práci slouží pouze jako korekce na reálný separační prostor, nezakreslený min-max škálováním. Druhý parametr, který nebyl v práci Schure a Davise zohledněn, byl případný vliv dimenzionalit vzorku.

Cílem této části disertační práce je pomocí modelů obsahujících 2500 bodů a reálných GC×GC separací nalézt deskriptor, který by nejlépe vystihoval současně globální a lokální hledisko ortogonality, a to jak u nízko- tak i vysokodimenzionálních vzorků.

## 10.2 Metody

### 10.2.1 Syntetické chromatogramy

Syntetické chromatogramy jsou v R-software vytvořeny v tak, že datové body vyplňují separační prostor o rozměrech 89-3000 s × 1-4 čtyřmi různými způsoby a tvoří 4 různé třídy modelů:

1. datové body vyplňují celý separační prostor a mají pravidelné rovnoměrné rozdělení (každý bod je umístěn přesně uprostřed svého binu)
2. datové body vyplňují celý separační prostor a mají náhodné rovnoměrné rozdělení
3. datové body tvoří shluky v prázdném separačním prostoru, body uvnitř shluků mají náhodné rovnoměrné rozdělení
4. datové body zcela vyplňují separační prostor a mají náhodné rovnoměrné rozdělení, část z bodů tvoří shluky

Každá třída chromatogramů se skládá z několika typů modelů (Tab. 5-8, S3, S4). Každý typ modelů s náhodným rozdělením bodů (tj. vyjma 1. typu) je generován stejným algoritmem 50×, aby bylo možné pomocí směrodatné odchylky ověřit validitu daného typu modelu. Algoritmy pro generování syntetických chromatogramů jsou obdobné těm, které byly použity v kapitole 9. Nejmenší vzdálenost mezi body v normalizovaném retenčním prostoru je 0,001.

### 10.2.2 Reálné chromatogramy

K ověření výsledků syntetických chromatogramů je použito reálných GC×GC separací močí, plasmy a nafty. Mrtvé časy jsou určeny pomocí software Column calculator a Chroma-TOF (v. 4.51.6.0, LECO, St. Joseph, Michigan, USA), a to na základě rozměrů použitých kolon, průtoku helia a teplotního programu). Mrtvý čas u syntetických chromatogramů je odvozen na základě reálných separací metabolomických vzorků, připravených ethoximací a trimethylsilylací. Pro přípravu moči se využívá kyselá extrakce do ethylacetátu [116], u plasmy deproteinace pomocí methanolu, odfoukání pod dusíkem, ethoximace a trimethylsilylace [117]. Detailní postup přípravy vzorků i jejich chromatogramy jsou uvedeny v kapitole 11.

### 10.2.3 Výpočet ortogonality

Retenční koordináty pík jsou škálovány pomocí WOSEL. Ortogonalita se stejně jako v kapitole 9 počítá několika různými přístupy vybranými tak, aby z každé skupiny přístupů o stejném principu byl vybrán právě jeden: Pearsonova korelace (PC) [34], podmíněná entropie (CE) [35], konvexní obálka (CH) [36], modelování (MA) [42], geometrické pokrytí separačního povrchu (SC) [38], metoda nejbližšího souseda (NN) [40] a hvězdičkové rovnice (AE) [41]. Principy jednotlivých přístupů a rovnice použité k výpočtu jsou uvedeny v úvodní části této disertační práce. Výpočty se provádějí v R-software (PC, CH, AE, MA a SC) a MATLAB (CE, NN, MATLAB 2018a, Natick, Massachusetts, USA).

## 10.3 Výsledky a diskuse

Cílem práce je nalezení „univerzální“ deskriptor ortogonality, který by byl jasnou volbou při optimalizaci 2D systému pro analýzu jak vysoko- tak i nízkodimenzionálních vzorků. Každý z deskriptorů, které jsou v této práci porovnávány, měří jinou vlastnost chromatogramu: korelaci mezi retenčními časy první a druhé dimenze [34], poměr počtu obsazených binů k prázdným [38] a jejich kombinaci [42]; poměr plochy obsazené píky k celkové využitelné ploše chromatogramu [36], průměrnou vzdálenost nejbližších sousedů [40], distribuci bodů podél čtyř os protínajících separační prostor [41] a míru shody distribuce bodů predikované na základě retenčních časů první dimenze se skutečnou distribucí bodů v dimenzi druhé [35]. Je tedy velmi těžké posoudit, který z deskriptorů je nejlepší.

Porovnání deskriptorů se provádí ve třech v následujících krocích:

1. Ověření schopnosti deskriptorů pokrýt celý rozsah ortogonalit od 0 do 100 %, a tím i jejich vhodnosti pro určení nejlepší a nejhorší separace; vyloučení deskriptorů, které toto kritérium nesplní.
2. Nalezení nejvhodnějšího deskriptoru pro současné určení globální i lokální ortogonalit jak nízko- tak i vysokodimenzionálních vzorků.
3. Stanovení vlivu cílené změny dimenzionality vzorku na optimalizaci analytického systému pomocí ortogonalit.

Protože se při vyhodnocování optimální separace 2D chromatografických systémů klade důraz nejen na jeho výslednou ortogonalitu, ale i na vzhled chromatogramu (viz kapitola 9), jsou číselné výsledky doplněny obrázky znázorňujícími disperzi retenčních koordinát v původním 2D separačním prostoru (Tab. 5-8, S3, S4, Obr. 37, 39-43). Díky použití WOSEL pro škálování je normalizovaný retenční prostor obdobný původnímu a není nutné je zobrazovat společně (viz kapitola 9).



Druhým způsobem prezentace výsledků použitým v této kapitole jsou srovnávací grafy (Obr. 40-42). Protože však každý z přístupů poskytuje pro konkrétní chromatogram jinou číselnou hodnotu a pro chromatogramy určitého typu jiný rozsah hodnot, je před konstrukcí srovnávacích grafů nutno hodnotu ortogonalitě obdrženou pro určitý chromatogram vydělit průměrnou hodnotou ortogonalitě počítanou stejným přístupem pro chromatogramy v daném porovnání.

Použití metod vícerozměrné statistické analýzy, doporučované autory předchozí práce [121], umožňuje pouze porovnání číselných hodnot, bez ohledu na vzhled chromatogramů. Dvojný graf PCA jedné skupiny modelů obsahující 4 různě umístěné shluky (Obr. 37) ukazuje pouze, že při této kombinaci proměnných do PC 1 a PC 2 (Obr. 38) CH a AE monitorují stejnou vlastnost chromatogramů, zcela odlišnou od HN. Na PCA není patrné žádné dělení modelů do skupin. Hlavní komponenty jsou tvořeny kombinací výsledků jednotlivých přístupů, kdy PC1 i PC2 je tvořena kombinací všech 5 přístupů a tím, že v PC1 je obsažen především AN, SC, HN a v PC2 AE, HN a CH. Nelze tedy říci, který z deskriptorů je pro popis pokrytí 2D prostoru nejlepší. Ze vzhledu chromatogramů přidaných na obvod Obr. 37 lze usuzovat, že SC ukazuje míru překrytí a CH a AE naopak rozptyl shluků.

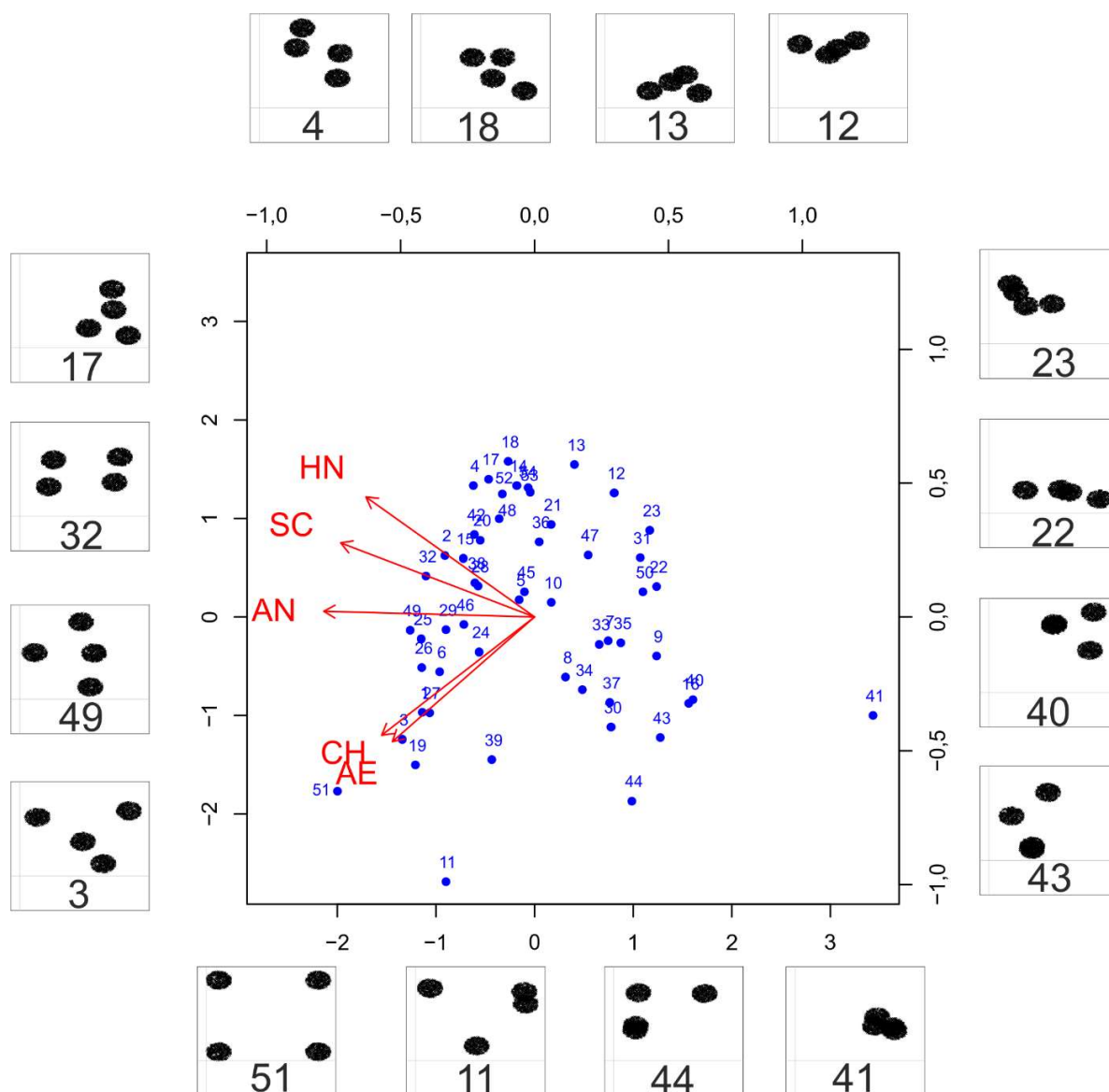
### 10.3.1 Rozsah deskriptorů

V této části kapitoly se vychází z předpokladu, že správný deskriptor je schopen pokrýt celý rozsah hodnot ortogonalitě od 0 do 100 % a že také umí odlišit průměrné separace od extrémních. Dva z deskriptorů (AN a HN), oba z přístupu NN, však nabývají zcela jiných hodnot (viz Tab. 3, kapitola 9). Pro modely s 2500 body jsou to hodnoty od 0,002 do 0,011, zatímco pro modely se 100 body je rozsah od 0,010 do 0,060 (Tab. S2). Je tedy zřejmé, že hodnoty AN i HN jsou silně závislé na počtu datových bodů, což vyplývá z výpočtů. Čím více bodů je obsaženo ve stejně velkém separačním prostoru, tím blíže jsou u sebe a průměrné vzdálenosti k nejbližším sousedům jsou menší. To je velkou nevýhodou přístupu NN, která brání jeho použití při optimalizaci systémů pomocí reálných vzorků (u kterých je např. v GC×GC při použití jiných kolonových uspořádání detekován jiný počet potenciálních metabolitů, tzv. features) a také jejich implementaci do kombinovaných deskriptorů [121]. Aby došlo k jejich odstranění, hodnoty AN a NN se přepočítávají na % pomocí v této práci nově vyvinutého vztahu (Rov. 12):

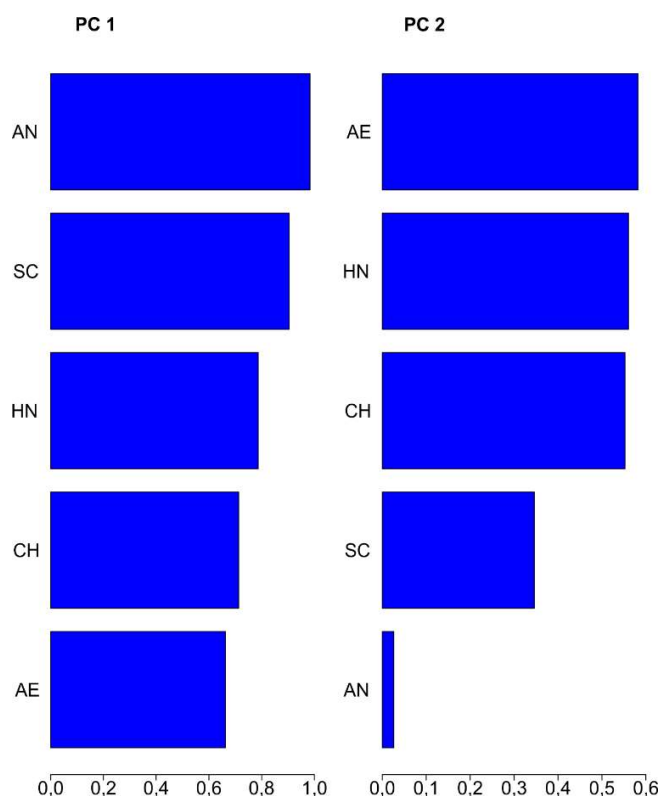
$$O = \frac{100 \% \cdot \bar{x} \cdot (\sqrt{\sum P} - 1)}{0,64} \quad (\text{Rov. 12}),$$

kde  $\bar{x}$  je AN nebo HN,  $\sum P$  je počet píků v separaci,  $(\sqrt{\sum P} - 1)$  je počet vzdáleností mezi nejbližšími sousedy v jedné dimenzi, 0,64 je maximální pokrytí separačního povrchu 2500 náhodnými veličinami

(píky) vyplývající z Poissonova rozdělení pravděpodobnosti [39]. Pro nižší počet píků může být hodnota 0,64 nahrazena hodnotou 0,63.



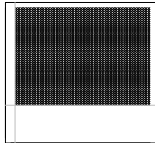
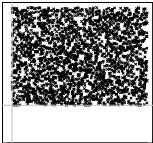
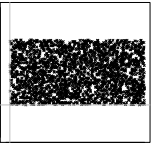
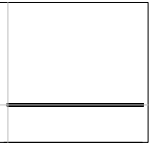
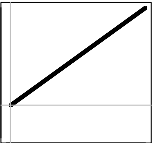
Obr. 37: Ukázka aplikace metod vícerozměrné statistické analýzy (dvojný graf PCA) na porovnání vybraných deskriptorů ortogonality u modelů se čtyřmi shluky. Významnost první komponenty je 67,06 %, druhé 21,56 %.



Obr. 38: Složení první a druhé hlavní komponenty PCA modelu na grafu uvedeném na Obr. 37

Po přepočtu AN a HN na % se společně se porovnává schopnost všech deskriptorů pokrýt celý rozsah ortogonalit od 0 do 100 %, a to pomocí 5 různých typů syntetických chromatogramů, vyobrazených v Tab. 5. Schopnost dosáhnout horní hodnoty rozsahu se testuje pomocí 2 typů chromatogramů, kde data vyplňují celý separační prostor. U typu A jsou datové body jsou umístěny uprostřed každého binu, u typu B mají data náhodné rozdělení. Přístupy AN, HN, MA a SC poskytují pro chromatogramy typu A hodnoty vyšší než 100 %. Je to proto, že tyto přístupy mají ve vzorcích pro výpočet ortogonalit inkorporovanou korekci na míru pravděpodobnosti toho, že u reálných jevů píky samy od sebe objeví vždy přesně uprostřed binu. Jak už bylo zmíněno dříve, dochází k tomu pouze u 64 % z 2500 píků. Proto je výsledná hodnota ve vzorci vydělena konstantou 0,64. Pokud jsou píky uměle umístěny do středu každého z binů, musí být výsledná hodnota ortogonalit vyšší než 100 %, jako je tomu v těchto 4 případech. Ostatní přístupy pravděpodobnost náhodných jevů nezohledňují.

Tab. 5: Výsledky ortogonality [%] ( $\pm 2$  SD pro modely s náhodným rozdělením) poskytnuté jednotlivými přístupy výpočtu pro modely pro určení rozsahu deskriptorů.

Typ	A	B	C	D	E	
Vzhled						
Rozdělení	Pravidelné	Náhodné	Náhodné	Pravidelné	Pravidelné	
Přístup	PC	100,0	98,8 $\pm$ 2,0	98,6 $\pm$ 2,2	96,5	0
	CE	100,0	85,8 $\pm$ 0,6	89,2 $\pm$ 0,8	-	0
	AN	156,3	99,9 $\pm$ 1,2	81,6 $\pm$ 1,1	3,1	4,4
	HN	156,3	67,3 $\pm$ 3,8	54,6 $\pm$ 3,6	3,1	4,4
	CH	96,0	99,2 $\pm$ 0,4	66,1 $\pm$ 0,3	0	0
	AE	95,9	96,3 $\pm$ 2,5	70,8 $\pm$ 0,8	0	0
	MA	158,7	100,3 $\pm$ 1,9	82,6 $\pm$ 2,0	3,1	0
	SC	158,7	100,3 $\pm$ 1,9	82,7 $\pm$ 2,0	3,1	3,2

Absence poklesu ortogonality mezi modely typu A a B pomocí CH není překvapivá, protože CH přináší informace o rozsahu využitého prostoru, ne o distribuci píků. To je patrné i z toho, že když jsou datové body u modelů typu A umístěny uprostřed binu a nezasahují i na jejich vnější okraje jako u modelů typu B, je ortogonalita separace A nižší o 4 % (v každé dimenzi je vyžito o  $2 \times 0,5$  binu méně. Proto má separační prostor rozsah 2401 místo 2500 binů, což přesně odpovídá 96 %). AE překvapivě nedosáhl hodnoty 100 % pro modely typu A ani B, neukázal také pokles hodnot mezi modely s náhodným a pravidelným rozdělením. V případě AE to však oproti CH nesouvisí s rozsahem využitého separačního prostoru, protože hodnoty získané pro modely typu A a B jsou více méně stejné. Z toho je zřejmé, že AE nepatří mezi deskriptory globální ortogonality.

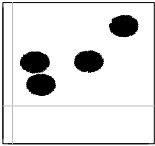
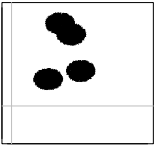
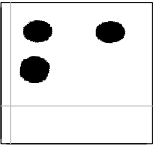
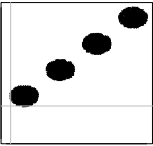
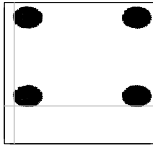
Nejvyšší pokles ortogonality mezi modely typu A a B je zaznamenán u HN. Je tomu tak proto, že HN, který se zaměřuje na monitorování krátkých vzdáleností mezi body, poukazuje na přítomnost velmi blízkých bodů. Nejbližší vzdálenost mezi body je v praxi dána chemometrickým rozlišením (protože od tohoto okamžiku jsou body vnímány jako 2 a ne jako jeden), u syntetických chromatogramů je tato vzdálenost nastavena analogicky těm reálným. Proto HN nedosahuje u modelů typu B hodnot 100 %.

Pro model typu C, kde je separační prostor náhodně vyplněný body o 1/3 menší než u modelu typu B, ukázaly všechny přístupy až na PC a CE nižší hodnoty než pro model B. Modely typu D a E imitují 1D separaci, ve které je druhá dimenze zcela nevyužitá. Mají podobu přímky rovnoběžné s první dimenzí (typ D) nebo úhlopříčky vycházející z počátku os (typ E). Ortogonalita těchto modelů by měla

být rovna 0. Hodnota ortogonality poskytnutá většinou přístupů se pohybuje v intervalu od 0 do 4,4, které jsou akceptovatelné, protože ortogonalita není ideální funkcí píkovej kapacity [39] (viz kapitola 11). CE neumožňuje spočítat ortogonalitu modelu typu D, protože u tohoto modelu není žádná nejistota v ose y. PC pro tento model poskytl ortogonalitu blízko 100 %.

Na základě Tab. 5 a Tab. 6 lze shrnout, že hodnoty ortogonality vypočítané pomocí PC a CE nereflktují využití separačního prostoru, a proto už nejsou podrobovány testování v dalších částech této práce, stejně jako MA, jehož je PC součástí. Přestože MA v některých případech správně rozlišuje dobré separace od těch horších, v případě, kdy lze body proložit přímkou, PC do výpočtu MA vnáší chybu (Tab. 6).

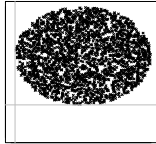
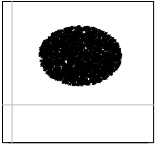
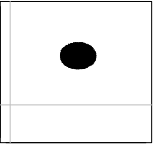
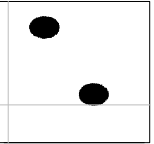
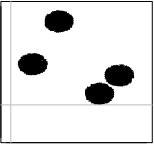
Tab. 6: Vliv korelace (R) na hodnoty ortogonality nízkodimenzionálních modelů typu E, spočítané pomocí PC, CE a MA

Model	2	37	44	51	52	
Vzhled						
Rozdělení	Náhodné	Náhodné	Náhodné	Náhodné	Náhodné	
Přístup	PC [%]	19,1	90,8	43,6	2,7	99,7
	CE [%]	72,3	81,3	87,5	58,0	95,4
	SC [%]	23,4	21,7	18,7	22,7	22,3
	R	0,81	0,09	0,56	0,97	0,00
	MA [%]	3,1	21,5	12,8	1,2	22,3

### 10.3.2 Nízkodimenzionální vzorky

Chromatogramy nízkodimenzionálních vzorků jsou charakteristické přítomností skupin(y) analytů v jinak prázdném 2D separačním prostoru. Látky v rámci každé skupiny mají podobné chromatografické chování, protože obsahují stejné funkční skupiny, jen jinak umístěné [45]. Příkladem takovýchto chromatogramů jsou GC×GC separace nafty [127], polycyklických aromatických uhlovodíků v environmentálních vzorcích [128], polychlorovaných bifenyly [129] nebo peptidů [121]. Syntetické chromatogramy připravené pro účely této práce obsahují jeden nebo více shluků v jinak prázdném separačním prostoru (Tab. 6, 7).

Tab. 7: Syntetické modelové chromatogramy nízkodimenzionálních vzorků a průměrná hodnota ortogonalita spočítaná jednotlivými přístupy pro každý typ modelů  $\pm 2SD$ .

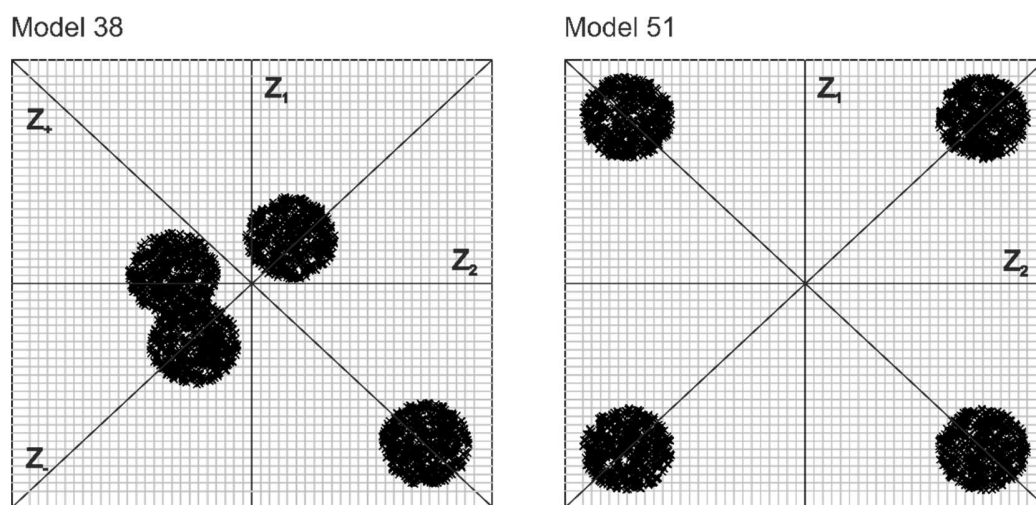
Typ	A	B	C	D	E	
Vzhled						
Rozdělení	Náhodné	Náhodné	Náhodné	Náhodné	Náhodné	
Přístup	AN	88,8 $\pm$ 0,5	54,1 $\pm$ 0,7	22,1 $\pm$ 0,1	25,7 $\pm$ 2,3	35,8 $\pm$ 4,3
	HN	59,7 $\pm$ 1,9	36,4 $\pm$ 2,6	14,8 $\pm$ 0,5	16,9 $\pm$ 1,6	23,0 $\pm$ 3,1
	CH	77,1 $\pm$ 0,1	27,8 $\pm$ 0,1	4,8 $\pm$ 0,0	11,4 $\pm$ 8,2	29,1 $\pm$ 26,8
	AE	78,3 $\pm$ 1,1	28,1 $\pm$ 0,7	4,9 $\pm$ 0,1	21,6 $\pm$ 24,1	43,6 $\pm$ 33,9
	SC	90,5 $\pm$ 0,9	44,8 $\pm$ 0,6	9,3 $\pm$ 0,1	11,7 $\pm$ 1,3	21,6 $\pm$ 3,7

Z výsledků uvedených v Tab. 7 je patrné, že všechny testované deskriptory jsou schopny zaznamenat změnu trendu syntetických chromatogramů. Pokles průměru shluku (modely typu A až C) je reflektován snížením ortogonalita u všech přístupů, bez ohledu na to, zda se jedná o deskriptor lokální či globální ortogonalita, protože dochází jak k poklesu využitého separačního prostoru, tak i k většímu nahromadění bodů (stejný počet bodů je umístěn na menší ploše). Navýšení počtu shluků (modely C až D) vede ze stejného důvodu k nárůstu ortogonalita u všech přístupů.

Hodnoty ortogonalita poskytnuté jednotlivými přístupy pro tentýž typ modelů se různí. Také jejich nárůst mezi jednotlivými typy modelů je rozdílný. Je to dáno odlišnými principy, na základě kterých se ortogonalita v rámci jednotlivých přístupů počítá [121]. Proto osamocená hodnota ortogonalita (např. 22,1 % pro model C pomocí AN) nemá pro ohodnocení chromatogramu význam, protože na základě čísla samotného není možné usuzovat na dobrou či špatnou separaci. Až po porovnání s hodnotami AN spočítanými pro ostatní modely je zřejmé, že se jedná o nepříliš dobrou separaci, protože hodnoty AN se mění od 89,3 do 22,0 %. Pro srovnání obdobnou hodnotu okolo 22 % poskytuje SC pro model E a AE pro model D. V případě SC se však hodnoty pohybují v rozmezí od 9,2 do 91,4 %, u AE od 4,8 do 79,4 %, takže tyto separace patří mezi „lepší“.

Nejvyšší variabilita výsledků mezi jednotlivými modely stejného typu je zaznamenána pro modely typu E obsahující 4 různě umístěné shluky, proto byly modely tohoto typu použity pro porovnání deskriptorů. V ideálním případě se shluky nepřekrývají (mají vysokou lokální ortogonalitu) a jsou umístěny co nejdále od sebe (mají vysokou globální ortogonalitu). Seznam výsledků získaných jednotlivými přístupy pro všechny modely typu E je uveden v Tab. S3.

Výsledky AE vykazují u modelů typu E největší rozsah hodnot (od 7,9 do 86,2 %), což by mohlo být důsledkem jeho vysoké citlivosti na změny pokrytí separačního prostoru. Obr. 39 však dokládá, že hodnoty ortogonality pomocí AE nevypovídají o pokrytí separačního prostoru, ale pouze o vzdálenosti shluků od os, zejména od jejich průsečíku. Proto modely se shluky umístěnými v blízkosti průsečíku os (model 38) mají vyšší ortogonalitu než modely s maximálně vzdálenými shluky umístěnými v rozích normalizovaného separačního prostoru (model 51). Tato vlastnost, AE společně s neschopností dosáhnout 100% ortogonality a nerefektování náhodné distribuce dat, je stěžejní pro vyloučení AE z dalšího testování.

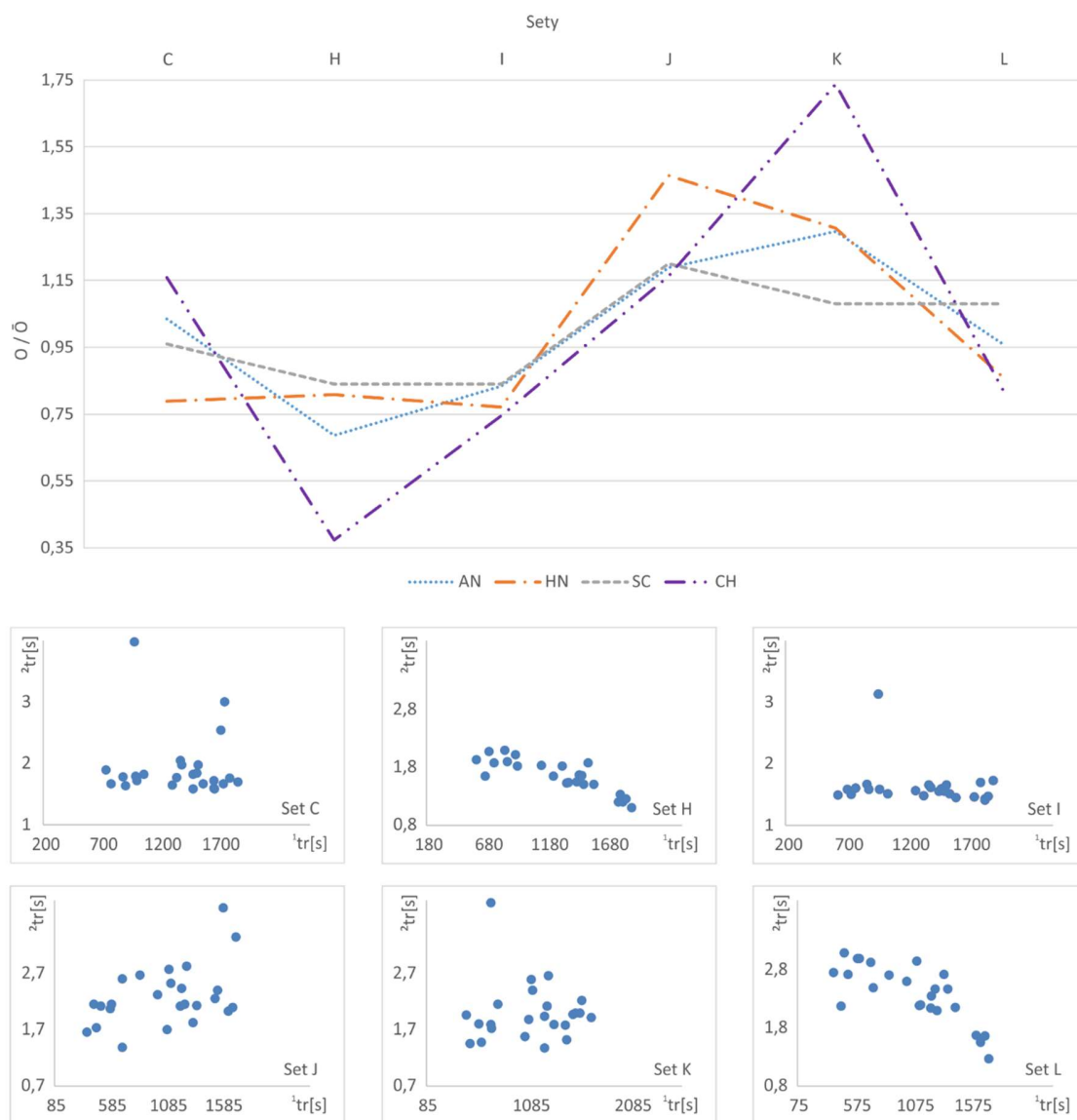


Model	Z [%]	Z+ [%]	Z1 [%]	Z2 [%]	O [%]
38	95,3	47,7	81,2	69,0	50,5
51	57,7	57,4	57,5	57,6	33,1

Obr. 39: Výpočet ortogonality dvou syntetických modelů typu E nízkodimenzionálních vzorků pomocí AE. Tabulka pod obrázkem uvádí distribuce bodů podél jednotlivých os ( $Z$ ,  $Z_+$ ,  $Z_1$ ,  $Z_2$ ) a výslednou ortogonalitu  $O$  [%] vypočítanou podle Rov. 7.

Také nízkodimenzionální modely potvrzují, že CH přináší pouze informaci o rozpětí dat v separačním prostoru, nikoli o vzdálenosti mezi jednotlivými píky [130], jak již bylo zmíněno v kapitole 10.3.1. Tab. S3 ukazuje, že CH je deskriptor globální ortogonality, protože pro model 51 poskytuje hodnotu blízkou maximu (98,5 %) bez ohledu na prázdné prostory mezi 4 shluky umístěnými v rozích separačního prostoru. Z toho plynou jeho 2 velké nevýhody – započítávání prázdného prostoru mezi píky do celkové ortogonality systému a vysoká citlivost k odlehlým hodnotám. Na reálných GC×GC separacích moči je manifestován vliv jediného distálního analytu na celkovou ortogonalitu systému

pomocí CH (Obr. 40). Přestože je chromatografický vzor setu H a I až na retenci jediné látky obdobný, ortogonalita pomocí CH je u setu I (17,7 %) hodnocena mnohem lépe než u setu H (8,8 %, kdy maximální dosažená hodnota pro set K je 41,1 %). Proto ani CH není vhodným deskriptorem pro popis ortogonalit nízkodimenzionálních vzorků.



Obr. 40: Optimalizace GCxGC systému pro analýzu 25 metabolitů v moči (převážně organických kyselin). Nahoře srovnávací graf s hodnotami ortogonalit normalizovanými průměrnou hodnotou ortogonalit pro jednotlivé přístupy, dole chromatografické vzory obdržené pro jednotlivá uspořádání kolon (set C až L).



Porovnání zbývajících tří deskriptorů, AN, HN a SC, je provedeno pomocí srovnávacího grafu zobrazeného na Obr. 41. AN se chová jako deskriptor globální ortogonality. Poskytuje nejvyšší hodnotu pro shluky umístěné v rozích separačního prostoru, které jsou od sebe vzdáleny nejvíce ze všech modelů (model 51,  $O = 41\%$ ), nejnižší pak pro čtyři překrývající se shluky (model 41,  $O = 28,5\%$ ). Maximální hodnota dosažená pro model 51 se na rozdíl od CH výrazně odlišuje od  $100\%$ , z čehož vyplývá, že AN do výsledné ortogonality započítává i prázdné prostory mezi shluky.

SC se chová jako deskriptor lokální ortogonality. Výsledné hodnoty ortogonality spočítané pomocí SC jsou pro jakékoli nepřekrývající se shluky skoro stejné (pohybují se v rozmezí od  $22,0$  do  $23,6\%$  v závislosti na rozmístění bodů náhodným algoritmem). SC tedy nereflektuje vzájemnou vzdálenost shluků. Pro překrývající se shluky vypočítaná hodnota klesá v závislosti na míře překryvu až k nejnižší hodnotě  $14,7\%$  pro model č. 41 (Obr. 41, Tab. S3).

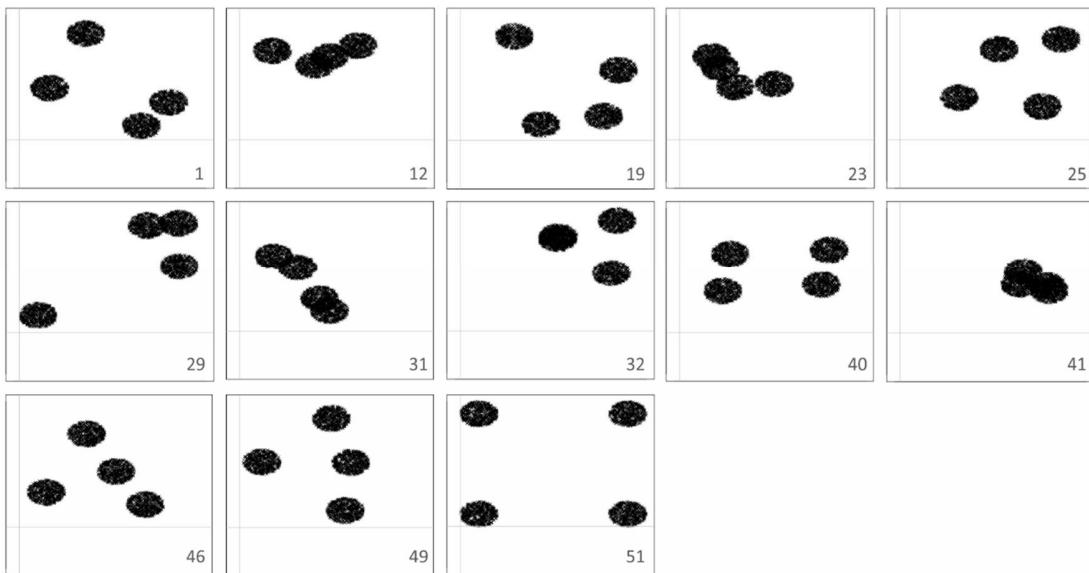
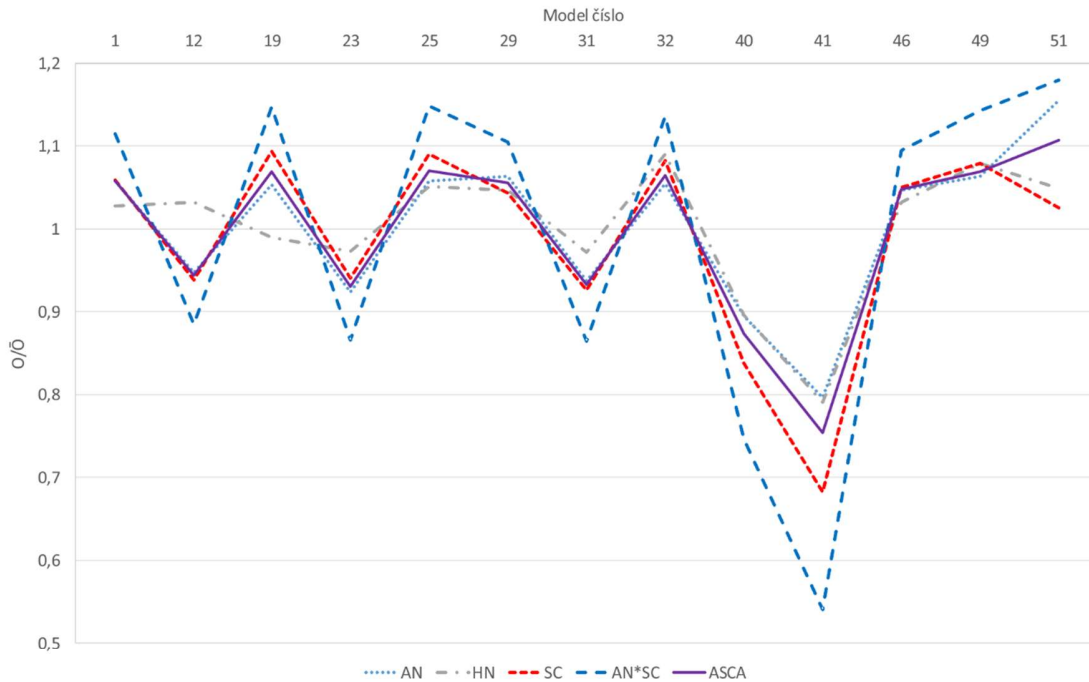
HN by se dle autorů tohoto přístupu měl také chovat jako deskriptor lokální ortogonality [40]. Nejvíce se překrývající shluky jsou správně ohodnoceny nejnižší ortogonalitou (model 41,  $O = 18,2\%$ ), ale z modelů 12 (blízké shluky, 3 z nich se překrývají) a 46 (4 jednotlivé shluky) už lepší separaci určit neumí (oba chromatogramy ohodnotil ortogonalitou  $23,7\%$ ). Je tomu tak nejspíš proto, že vysoký počet větších vzdáleností potlačil vliv vzdáleností menších.

Nejlepším lokálním deskriptorem pro nízkodimenzionální vzorky je proto SC, globálním pak AN. Pro každý chromatogram tak lze získat dvě různé hodnoty a při vyhodnocování nejlepších separačních podmínek si nakonec vybrat buď jeden nebo druhý aspekt ortogonality. Protože maximální využití separačního prostoru a minimální počet shluků je pro separaci stejně důležité, navrhli Schure a Davis vytvořit kombinovaný deskriptor vynásobením dvou nejlepšího lokálního a globálního deskriptoru [121]. V jejich práci jím byla dimenzionalita (Dim, jedna z metod založená na počítání obsazených binů [126]) a CH. Vynásobení AN a SC však v této disertační práci nepřináší žádný užitek, pouze zesiluje trend daný původními deskriptory (špatná separace se stává ještě horší, dobrá ještě lepší, Obr. 41). Benefit vynásobení hodnoty Dim hodnotou CH v práci Schurea a Davise byl dán nejspíš tím, že představoval korekci na min-max škálování. CH se totiž na rozdíl od Dim počítalo v původním, neškálovaném separačním prostoru.

Při použití WOSEL násobení deskriptorů ztrácí význam. Za předpokladu stejné důležitosti globální a lokální ortogonality se jako nejlepší způsob kombinace jeví výpočet průměru obou hodnot. Protože se však oba deskriptory v konkrétním souboru vzorků liší rozsahem hodnot, je výhodnější použít k zprůměrování hodnot poskytnutých SC a AN geometrický průměr. Hodnoty nového kombinovaného deskriptoru ASCA (z angl. „AN and SC average“) se pak spočítají podle Rov. 13.

$$O_{ASCA} = \sqrt{O_{AN} \cdot O_{SC}} \quad (\text{Rov. 13}),$$

$O_{ASCA}$  je hodnota ortogonality [%] spočítaná pomocí nového deskriptoru ASCA,  $O_{AN}$  je hodnota ortogonality spočítaná podle Rov. 6 a 12,  $O_{SC}$  je hodnota ortogonality spočítaná podle Rov. 4.

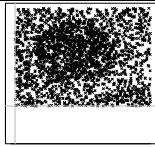
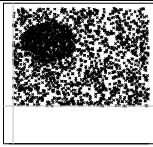
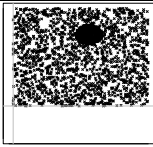
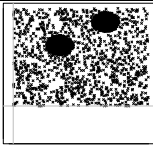
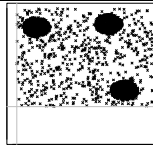


Obr. 41: Porovnání AN, HN a SC a kombinovaných deskriptorů AN\*SC a ASCA pomocí nízkodimenzionálních syntetických modelů typu E. Nahoře srovnávací graf s hodnotami ortogonality normalizovanými průměrnou hodnotou ortogonality pro jednotlivé přístupy, dole chromatogramy jednotlivých modelů.

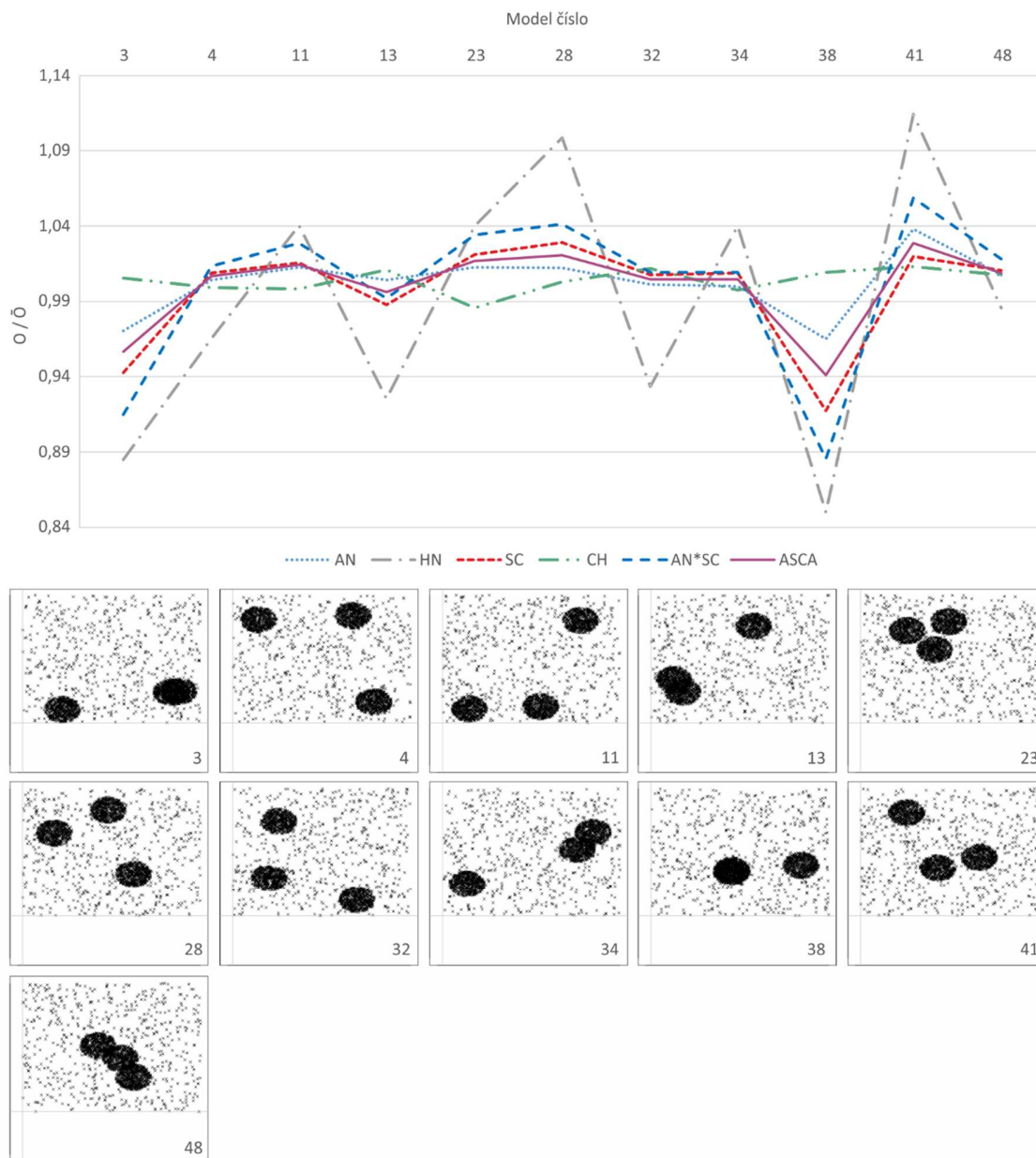
### 10.3.3 Vysokodimenzionální vzorky

Vysokodimenzionální vzorky nepředstavují tak častý analytický problém jako celá řada vzorků obsahující obrovská množství strukturně podobných látek (nafta, minerální oleje, esenciální oleje, cigaretový kouř, polychlorované bifenylly, pesticidy, methylestery mastných kyselin, steroidy aj.). Vysokodimenzionální vzorky se skládají z látek obsahujících celou řadu funkčních skupin, které různou měrou interagují se stacionární fází první a/nebo druhé dimenze. Na chromatogramech těchto vzorků nejsou přítomny izolované skupiny strukturně podobných látek, datové body jsou rozesety po celém chromatogramu [45]. S vysokodimenzionálními vzorky se setkáváme při analýze metabolomických vzorků [131], tabákových extraktů [132] nebo kávy [133]. Syntetické chromatogramy (Tab. 8) jsou vytvořeny tak, že celý jejich 2D separační prostor je vyplněn datovými body, do kterých jsou umístěny shluky dalších bodů. Celkový počet bodů je vždy 2500 a mají náhodné rovnoměrné rozdělení. Každý shluk obsahuje vždy ¼ z celkového počtu bodů. Nový shluk je vytvořen z bodů mimo shluky.

Tab. 8: Syntetické modelové chromatogramy vysokodimenzionálních vzorků a průměrná hodnota ortogonalita spočítaná jednotlivými přístupy pro každý typ modelů  $\pm 2SD$ .

Typ	A	B	C	D	E	
Vzhled						
Rozdělení	Náhodné	Náhodné	Náhodné	Náhodné	Náhodné	
Přístup	AN	98,1 $\pm$ 1,3	96,3 $\pm$ 0,9	92,8 $\pm$ 1,4	83,4 $\pm$ 1,7	71,1 $\pm$ 2,1
	HN	63,1 $\pm$ 4,0	58,7 $\pm$ 3,1	46,1 $\pm$ 1,9	34,6 $\pm$ 3,5	27,7 $\pm$ 3,8
	CH	99,0 $\pm$ 0,4	98,9 $\pm$ 0,5	98,9 $\pm$ 0,5	98,5 $\pm$ 0,7	97,3 $\pm$ 1,4
	AE	87,4 $\pm$ 4,6	86,6 $\pm$ 7,9	85,9 $\pm$ 6,1	70,5 $\pm$ 13,4	59,8 $\pm$ 23,2
	SC	96,0 $\pm$ 1,9	91,5 $\pm$ 1,8	85,5 $\pm$ 1,9	68,3 $\pm$ 1,6	47,8 $\pm$ 2,5

Všechny testované deskriptory s výjimkou CH jsou schopny odrážet změnu trendu syntetických modelů: pokles průměru shluku u modelů typu A až C i nárůst počtu shluků (a tím i pokles bodů mimo shluky) u modelů typu C až E vede k poklesu ortogonalita. CH soustavně ukazuje, že u všech typů modelů je celý separační prostor pokryt datovými body. U modelů typu E je datových bodů mimo shluky nejméně, proto jsou v některých případech okraje chromatogramů prázdné. Pak CH ukazuje malý pokles využitého separačního prostoru (Obr. 42). Pokud jsou prázdná místa uprostřed chromatogramu, pak je nereflkuje. AE stejně jako u nízkodimenzionálních modelů vykazuje vysokou variabilitu výsledků a ukazuje pouze vzdálenost shluků od (středu) os bez ohledu na vzájemnou vzdálenost shluků.



Obr. 42: Porovnání CH, AN, HN a SC a kombinovaných deskriptorů AN\*SC a ASCA pomocí vysokodimenzionálních syntetických modelů typu E. Nahoře srovnávací graf s hodnotami ortogonality normalizovanými průměrnou hodnotou ortogonality pro jednotlivé přístupy, dole chromatogramy jednotlivých modelů.

Zbylé tři deskriptory, AN, HN a SC, jsou testovány pomocí modelů typu E s nejvyšší variabilitou výsledků (Obr. 42, Tab. S4). HN má z nich nejvyšší variabilitu výsledků, což by mohlo svědčit pro nejvyšší citlivost deskriptoru. Ale není tomu tak. Změna hodnoty ortogonality není spojena se změnou vzájemné

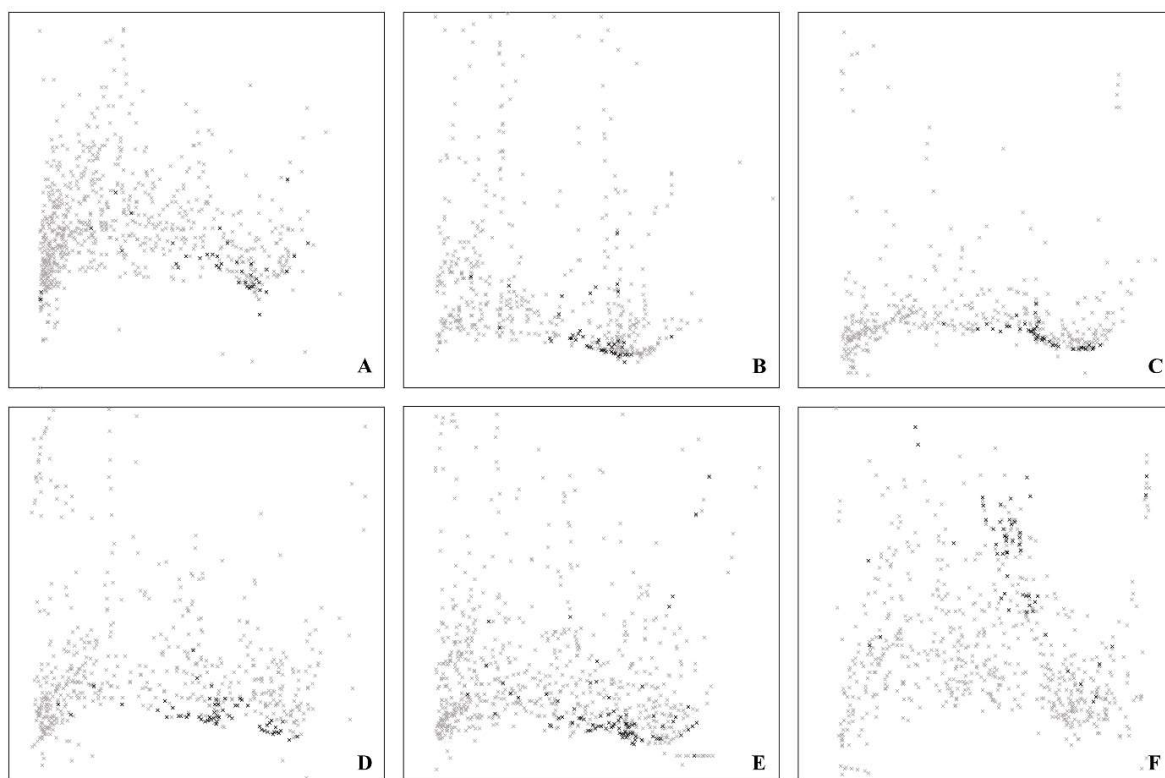
polohy shluků, protože vzdálené, blízké i překrývající se shluky jsou prostřednictvím HN ohodnoceny stejnou ortogonalitou (model 11, 23 a 34;  $O = 28,8\%$ ). Oproti tomu modely stejného vzhledu (28 a 32) mají při použití HN algoritmu ortogonalitu rozdílnou (30,4 vs. 25,8 % při rozsahu hodnot od 22,1 do 30,9 %). Tento jev je nejspíše způsoben zprůměrováním velkého počtu vzdáleností a faktem že každé prodloužení vzdálenosti mezi dvěma body je vyváženo zkrácením vzdálenosti mezi jinými dvěma body.

Ani přístup AN není schopen zachytit změnu vzájemné polohy shluků. Např. model 41 byl ohodnocen jako nejlepší ( $O = 73,9\%$ ), přestože u modelu 4 ( $O = 71,5\%$ ) jsou vzdálenosti mezi shluky mnohem větší. Na druhou stranu překrývající se shluky (model 13) mají ortogonalitu stejnou ( $O = 71,5\%$ ) jako vzdálené (model 4). Je tomu tak nejspíš proto, že AN je deskriptor globální ortogonalita a jako takový ukazuje, že celý separační prostor je pokryt datovými body. Proto jsou jeho hodnoty pro všechny modely víceméně neměnné (od 68,7 do 73,9 %).

Lokální ortogonalitu nejlépe dokládá SC, neboť všechny překrývající se shluky mají při použití tohoto algoritmu sníženou ortogonalitou (např. model 3,  $O = 45,1\%$ ). Nejnižších hodnot dosahuje model 38 ( $O = 43,9\%$ ), jehož shluky se zcela překrývají. Jsou-li shluky pouze v těsné blízkosti, k poklesu ortogonalita už nedochází (model 48,  $O = 71,7\%$ ). Její hodnota je obdobná jako u nepřekrývajících se shluků (model 41), jak je typické pro všechny deskriptory lokální ortogonalita.

Z analýzy syntetických modelů vysokodimenzionálních vzorků vyplývá, že globální ortogonalita není v tomto případě důležitým parametrem. Ukazuje pouze to, že je celý separační prostor vyplněn body a nabývá víceméně konstantních hodnot. Nelze ji nijak optimalizovat. Lokální ortogonalitu oproti tomu optimalizovat lze, a to pomocí SC. Kombinovaný deskriptor ASCA nemá příliš velký význam. Ale vzhledem k tomu, že AN je v podstatě konstanta, tak jeho průměrování s SC nevnáší do vyhodnocení nejlepší separace žádnou chybu. Přístup ASCA lze tedy použít i pro optimalizaci vysokodimenzionálních vzorků.

Stejné závěry vyplývají i z vyhodnocení reálných GC×GC chromatogramů vzorku lidské plasmy, analyzované pomocí 6 různých kolonových uspořádání (Obr. 43). AN ani HN neshledávají rozdíl mezi separací vzorku za pomoci kolonového uspořádání A a B. Ortogonalita spočítaná pomocí SC je v souladu se vzhledem chromatogramů, protože jako nejlepší hodnotí kolonové uspořádání F, nejhorší C. Při vyhodnocování optimálních separačních podmínek pomocí reálných vzorků je nutné vzít v úvahu, že vzhled chromatogramu, který je při vyhodnocování vnímán lidským okem, je silně ovlivněn (a zkrácen) počtem nalezených píků potenciálních metabolitů, které kromě analytů zahrnují i celou řadu artefaktů.



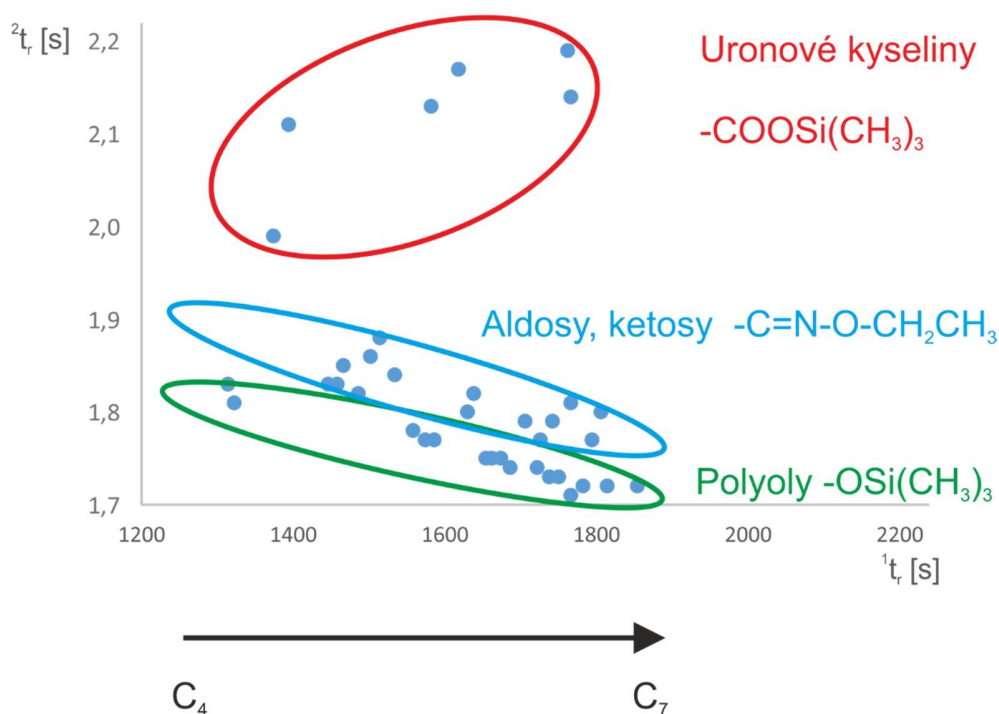
Deskriptor	Kolonové uspořádání					
	A	B	C	D	E	F
AN	57,6 (35,1)	58,0 (26,6)	50,6 (13,2)	62,2 (22,8)	65,4 (38,0)	66,3 (47,8)
HN	30,6 (15,6)	29,1 (10,0)	23,8 (8,0)	32,7 (12,4)	34,3 (15,6)	40,1 (18,1)
SC	45,2 (41,4)	49,3 (31,5)	39,5 (16,8)	51,6 (27,0)	52,1 (36,1)	58,7 (43,2)
ASCA	51,4 (38,3)	53,7 (29,1)	45,1 (15,0)	56,9 (24,9)	58,8 (37,1)	62,5 (45,5)

Obr. 43: GCxGC analýza plasmy pomocí šesti různých kolonových uspořádání (A – F). TIC šedě, EIC 217 černě. Tabulka pod obrázkem uvádí ortogonalitu [%] spočítanou pomocí jednotlivých deskriptorů pro TIC (EIC).

#### 10.3.4 Požadovaná dimenzionalita

Biologické vzorky obsahují tisíce látek mnoha funkčních skupin. Jejich chromatogramy ani zdaleka nejsou strukturované. V necílené metabolomice jsou všechny látky přítomné ve vzorku důležité. Vzorek je posuzován jako vysokodimenzionální. Oproti tomu v cílené metabolomice nebo klinické diagnostice je předmětem zájmu pouze určitá skupina látek. Všechny ostatní látky přítomné ve vzorku nejsou předmětem zájmu. Na vzorek původně vysokodimenzionální je nyní nahlíženo jako na nízkodimenzionální.

Vysokodimenzionální vzorky lze na nízkodimenzionální převést při přípravě vzorku (např. extrakcí vybraných látek) nebo použitím specifického detektoru (např. detektor elektronového záchytu, plamenový ionizační detektor se solí alkalického kovu). V případě detekce hmotnostní spektrometrií lze snížení dimenzionality docílit aplikací specifického  $m/z$  filtru (a vytvořením extrahovaného iontového chromatogramu, EIC), jako je tomu u separací vzorku plasmu pomocí šesti různých kolonových uspořádání (Obr. 43). Při záznamu celkového iontového proudu (TIC, na Obr. 43 znázorněn šedě) zachycujícího všechny píky potenciálních metabolitů chromatogram postrádá vnitřní strukturu, jak je typické pro vysokodimenzionální vzorky. Po aplikaci filtru pro  $m/z$  217 (černě) dochází ke snížení dimenzionality za účelem analýzy látek cukerné povahy. Vzniká strukturovaný chromatogram, kde na ose  $x$  jsou analyty řazeny dle délky uhlíkového řetězce, na ose  $y$  podle funkčních skupin (polyoly, cukry, uronové kyseliny, Obr. 44). V EIC (Obr. 43) se AN stává informativním, referuje o rozsahu využitého separačního prostoru, zatímco SC je indikátorem míry nahromadění dat v obou datových souborech (TIC i EIC). Dle SC i ASCA je nejlepším kolonovým uspořádáním pro necílenou analýzu plasmu (vysokodimenzionální vzorek) set F, pro cílenou analýzu látek cukerné povahy (nízkodimenzionální vzorek) dle SC set F nebo A, dle AN set F, dle ASCA set F (Obr. 43).



Obr. 44: Strukturovaný chromatogram, který vzniká po aplikaci filtru  $m/z$  217 z TIC vzorku plasmu změřeného pomocí kolonového uspořádání D uvedeného na Obr. 43.

## 10.4 Závěr

Optimalizace vícerozměrných chromatografických systémů pro analýzu nízkodimenzionálních vzorků představuje běžný analytický problém. Jedním z příkladů je cílená metabolomická analýza. K optimalizaci se nejlépe hodí nově vyvinutý deskriptor ASCA, který umožňuje optimalizovat oba aspekty ortogonalit, maximální pokrytí separačního prostoru a minimální shlukování dat, v jedné veličině. ASCA v sobě formou geometrického průměru kombinuje hodnotu ortogonalit získanou SC algoritmem, který nejlépe ze všech testovaných přístupů referuje o lokální ortogonalitě, a AN, jehož hodnoty odrážejí globální ortogonalitu.

Analýza vícedimenzionálních vzorků je méně častá, setkáváme se s ní např. u necílené metabolomiky. Jelikož datové body pokrývají celý separační prostor, optimalizace globálního aspektu ortogonalit nemá význam. AN se chová víceméně jako konstanta. Lokální ortogonalitu nejlépe dokládá SC algoritmus, jehož nejnižší hodnoty svědčí pro maximální shlukování dat. Pokud je přístup SC zkombinován do ASCA jako u nízkodimenzionálních dat, výsledky optimalizace to neovlivní, neboť se do výsledné hodnoty připočítává konstanta. Tak lze ASCA použít i pro optimalizaci vysokodimenzionálních vzorků.

Dimenzionalita reálných vzorků je obrovská, pohybuje se v řádu tisíců. Ne všechny vlastnosti analytů jsou však důležité k dosažení cíle analýzy. Proto lze ve většině případů dimenzionalitu vzorku snížit, vzorek převést na nízkodimenzionální a analytický systém pak plně optimalizovat na oba aspekty ortogonalit.



## 11 Nalezení nejvíce ortogonální kombinace kolon pro GC×GC analýzu metabolomu moči

Počet píků, které je chromatografický systém schopen rozlišit během jedné separace, je kompromisem mezi délkou analýzy a detekcí nejmenšího píku (tzv. minimální detekovatelná koncentrace) [134]. Nejčastěji je vyjadřován pomocí píkové kapacity ( $n_c$ ), kterou lze v plynové kapilární chromatografii spočítat podle rovnice (Rov. 14, [135]):

$$n_c = 1 + \frac{\sqrt{\bar{n}}}{4} \cdot \log \frac{t_r}{t_0} \quad (\text{Rov. 14}),$$

kde  $\bar{n}$  je průměrný počet teoretických pater kolony (díky měnícímu se kapacitnímu faktoru není konstantní),  $t_r$  je retenční čas,  $t_0$  je mrtvý čas.

Komplexní vzorky obsahují tisíce látek a cílem analýzy je jich detekovat co nejvíce (maximalizovat píkovou kapacitu). V jednodimenzionální plynové chromatografii toho lze docílit použitím delší kolony. Tím se však analýza mnohonásobně prodlouží (např. k zdvojnásobení píkové kapacity je třeba osmkrát delší analýza) [134]. Oproti tomu spojení s druhou kolonou jiné selektivity při kompletních (angl. „comprehensive“) dvojdimenzionálních analýzách vede k navýšení píkové kapacity  $n_c(\text{GC} \times \text{GC})$  při zachování času analýzy (Rov. 15, [136]):

$$n_c(\text{GC} \times \text{GC}) = K \cdot n_c(D1) \cdot n_c(D2) \quad (\text{Rov. 15}),$$

kde  $n_c(D1)$  je píková kapacita první a  $n_c(D2)$  druhé dimenze,  $K$  je konstanta vyplývající z chemismu stacionárních fází kolon.

Parametry  $n_c(D1)$  a  $n_c(D2)$  v 2D analytickém systému se odlišují od těch, které lze při stejných rozměrech kolon získat v 1D systému (kde jsou měřeny pro každou z kolon zvlášť). Je tomu tak proto, že při převádění analytů z první a druhé dimenze dochází k rozšíření jejich píků [134]. Kolona druhé dimenze má také vliv na tvar píků v první dimenzi. Je-li průměr druhé kolony menší než první, dochází ke zvýšení tlaku v koloně první dimenze a tím k rozšiřování píků první dimenze (fázová rovnováha se posouvá směrem ke stacionární fázi). Pokud se v druhé dimenzi používá kolona stejného průměru jako v první, druhá kolona separaci v první koloně téměř neovlivní. I přesto je výsledná píková kapacita GC×GC systému rovna přibližně polovině násobku píkové kapacity daných kolon používaných jednotlivě [137].

Dalším faktorem, který je třeba zvážit, je rozdíl v selektivitě stacionární fáze první a druhé dimenze. V rov. 15 je zastoupen konstantou  $K$  neznámé hodnoty. Pokud se k analýze vzorku použijí kolony zcela jiné (ortogonální), její hodnota je rovna jedné. Tento jev v praxi nenastává, neboť separační

mechanismy první dimenze se spolu se separačními mechanismy druhé dimenze vždy do jisté míry shodují. Typickým příkladem je použití polymethylsiloxanové kolony s 5% fenylu v první dimenzi a polymethylsiloxanové kolony s 50 % fenylu v druhé dimenzi, které je dodnes v GC×GC nejběžnější. Díky částečné shodě v chemismu stacionárních fází je příspěvek druhé kolony k celkové separaci nižší. Míru tohoto poklesu nelze přesně vyčíslit. Závisí jak na kombinaci stacionárních fází, tak na vlastnostech analytů vzorku. Proto píková kapacita není vhodným nástrojem k popisu účinnosti separace GC×GC systému [134]. V této práci je k tomuto účelu využíváno ortogonalita.

## 11.1 Kolony v GC×GC

Výběr kolon pro GC×GC se řídí snahou separovat co nejvyšší počet komponent vzorku. Proto se kombinují kolony co nejvíce rozdílné (ortogonální), lišící se separačními mechanismy. Tzv. normální kolonové uspořádání je tvořeno nepolární kolonou v první dimenzi, která separuje látky podle disperzních sil (eluční pořadí pak koresponduje s bodem varu látek), a polární (více selektivní) kolonou v dimenzi druhé, jež separuje koelující látky první dimenze podle polarit. Například při analýze nafty první nepolární dimenze umožní separaci alkanů a cykloalkanů, zatímco areny jsou díky  $\pi$ - $\pi$  interakcím zadržovány druhou dimenzí. V tzv. reverzním kolonovém uspořádání je tomu naopak [138].

Výběr kombinace kolon pro GC×GC analýzu se dodnes často řídí metodou pokus-omyl [138], popřípadě se přejímají kolonová uspořádání uvedená v literatuře [139]. Někteří dodavatelé kolon nabízejí balíčky pro GC×GC (angl. column selectivity kit), obsahující kolony různého průměru a selektivity, a také průvodce výběrem kolon (angl. „GC×GC column combination guide“). Zmiňovaný průvodce však není určen pro vzorky k metabolické analýze.

Obecně lze říci, že chemické složení stacionární fáze je pro vznik interakcí s analytem rozhodující. Za ideální stacionární fázi je považována taková, která umožní separaci analytů s využitím minimální píkové kapacity v co nejkratším čase [134]. Tloušťka vrstvy stacionární fáze má vliv na tvar píku (a s ním související parametry) a na zatížitelnost kolony (angl. „mass loadability“). Čím složitější matrice a větší koncentrační rozsah analytů, tím se volí širší kolona a silnější vrstva stacionární fáze [140]. Při volbě optimální délky druhé kolony je třeba mít na zřeteli, že delší kolona má vyšší mrtvý čas, čímž se využitelný separační prostor druhé dimenze zmenšuje. Na druhou stranu kratší D2 kolona má méně teoretických pater a tím i nižší píkovou kapacitu [137].

Tradičním kolonovým uspořádáním používaným v GC×GC je spojení nepolární 100% polymethylsiloxanové kolony se středně polární polymethylsiloxanovou kolonou s 50 % fenylu v dimenzi druhé [141]. Časté je také použití polymethylsiloxanové kolony s 5 % fenylu, běžné v 1D GC

[142], pro první dimenzi GC×GC. Při tomto přístupu zůstane zachováno obvyklé eluční pořadí, umožňující identifikaci analytů pomocí retenčních indexů v tradičních MS knihovnách (např. GOLM, NIST). Frakcionační kapacita se navýší připojením druhé kolony jiné selektivity, často s 50 % fenylu [29]. Oproti tomu reverzní kolonové uspořádání (polymethylsiloxan s 50 / 5 % fenylu) přináší zhoršení celkové separace, ale díky lepšímu tvaru píku a tím i vyšší citlivosti může být výhodné v cílené analýze, jako je tomu u stanovení sterolů v moči [143] či metabolitů v séru [140].

Souhrn kolonových uspořádání, které se používají pro analýzu metabolomických vzorků derivatizovaných trimethylsilylací, uvádí Tab. 9. Ve většině případů je použito normální uspořádání s nepolární kolonou 30 m polymethylsiloxanovou kolonou s 5 % fenylu v první dimenzi s vnitřním průměrem 0,25 mm, tloušťkou stacionární fáze 0,25  $\mu\text{m}$  a polymethylsiloxanovou kolonou s 50 % fenylu (např. BPX-50, Rxi-17, DB-17, TRB-50) v dimenzi druhé. Délka kolony druhé dimenze se většinou pohybuje mezi 1 a 2 m. Vnitřní průměr a tloušťka vrstvy stacionární fáze kolony druhé dimenze se různí podle typu vzorku. Nejširší kolony se používají pro analýzu séra (Tab. 9).

Z Tab. 9 je zřejmé, že ve většině metabolomických aplikací převládá tradiční kolonové uspořádání. Cílem této části disertační práce je pomocí ortogonálního přístupu (deskriptoru ASCA) otestovat jiné druhy stacionárních fází a nalézt nejlepší kolonové uspořádání pro analýzu metabolomu moči a séra. U nejlepších kolonových uspořádání také určit vliv změny vnitřního průměru kolony druhé dimenze na ortogonalitu systému.

## 11.2 Chemikálie

V práci se používají následující chemikálie: 4-fenylbutyrát (4-PB), methyl- $\text{D}_3$ -malonát (MMAD),  $^{13}\text{C}_3$ -laktát (CLAC), methanol pro LC-MS, voda pro LC-MS, O-ethylhydroxylamin hydrochlorid (EtOx), pyridin, bis(trimethylsilyl)trifluoroacetamid + 1% chlortrimethylsilan (BSTFA), N-Trimethylsilyl-N-methyl trifluoroacetamid (MSTFA) a ethylacetát od Sigma-Aldrich; chlortrimethylsilan pro GC-MS (TMCS) od Merck; síran sodný bezvodý, chlorid sodný, koncentrovaná kyselina chlorovodíková a aceton od lach:ner,  $\text{D}_8$ -fenylalanin (PHED) od Cambridge Isotop Laboratories. Všechny použité chemikálie jsou v čistotě *p.a.* nebo vyšší.

Tab. 9: Kolonová uspořádání používaná v literatuře pro metabolickou GC×GC analýzu analytů séra (S), plasmy (P), moči (M), buněk (B), tkání/pletiv (T), sputa (H) derivatizovaných trimethylsilylací.

Kolony (D1/D2)	Rozměry kolon*	Materiál						Typ analýzy	Odkaz
		S	P	M	B	T	H		
SHM5MS/BPX-50	D1: 30 - 0,25 - 0,25 D2: 5,0 - 0,15 - 0,15	X	X	X	X			necílená	[29]
BPX50/BPX5	D1: 30 - 0,25 - 0,25 D2: 2,0 - 0,32 - 0,25	X						cílená	[140]
BPX50/BPX5	D1: 30 - 0,25 - 0,25 D2: 2,0 - 0,32 - 0,25					X		necílená	[144]
Rxi-5Sil MS/Rxi-17	D1: 30 - 0,25 - 0,25 D2: 1,0 - 0,10 - 0,10				X		X	necílená	[145]
Rxi-5Sil MS/Rxi-17	D1: 30 - 0,25 - 0,25 D2: 1,2 - 0,25 - 0,25				X			necílená	[146]
DB-5 MS/DB-17H	D1: 30 - 0,25 - 0,25 D2: 1,2 - 0,25 - 0,25					X		necílená	[147]
HP5/Rtx-200	D1: 28 - 0,32 - 0,25 D2: 1,75 - 0,25 - 0,25				X			necílená	[148]
SE52/OV1701	D1: 30 - 0,25 - 0,25 D2: 1,4 - 0,10 - 0,10			X				necílená cílená	[149]
DB-5MS/RTX200	D1: 30 - 0,25 - 0,25 D2: 1,5 - 0,18 - 0,25	X						necílená	[150]
ZB-5MS/TRB-50	D1: 30 - 0,25 - 0,25 D2: 2,0 - 0,25 - 0,25					X		necílená	[151]
DB-5MS/DB17 MS	D1: 30 - 0,25 - 0,25 D2: 1,79 - 0,15 - 0,15					X		necílená	[152]
Rxi-5SilMS/Rxi17	D1: 30 - 0,25 - 0,25 D2: 1,0 - 0,10 - 0,10				X			necílená	[153]
Rxi-5Sil MS/BPX50	D1: 15 - 0,25 - 0,25 D2: 1,1 - 0,15 - 0,15			X				necílená	[154]
RTX-5MS/RTX-2000MS	D1: 20 - 0,25 - 0,5 D2: 2,0 - 0,18 - 0,20				X			necílená	[155]
BPX50/BPX5	D1: 30 - 0,25 - 0,25 D2: 2,0 - 0,32 - 0,025			X				cílená	[156]
Rxi-5SilMS/BPX50	D1: 20 - 0,25 - 0,25 D2: 2,6 - 0,15 - 0,15			X				necílená	[157]
DB-5MS/DB-17MS	D1: 60 - 0,25 - 0,25 D2: 1,0 - 0,25 - 0,25	X						necílená	[158]
BPX50/Rtx200	D1: 30 - 0,25 - 0,25 D2: 1,0 - 0,18 - 0,10					X		necílená	[159]
DB-1/Rxi17	D1: 30 - 0,25 - 0,25 D2: 1,5 - 0,10 - 0,10			X				necílená	[160]

- \* D1 - první dimenze, D2 – druhá dimenze; rozměry kolon jsou ve formátu: délka kolony [m] - vnitřní průměr [mm] - tloušťka vrstvy stacionární fáze [μm]

## 11.3 Příprava vzorku

### 11.3.1 Moč

K stanovení ortogonalita kolonových uspořádání se používá směsné moči, získané smícháním močí zdravých dobrovolníků. K 1 ml moči ve skleněné zkumavce se přidávají interní standardy: 20  $\mu\text{l}$  4-PB (2,11  $\text{mmol.l}^{-1}$ ) a 10  $\mu\text{l}$  MMAD (1,04  $\text{mmol.l}^{-1}$ ), 0,03 g EtOx a důkladně se vše promíchává. Po 30 min stání při pokojové teplotě se přidává 100  $\mu\text{l}$  HCl (6  $\text{mol.l}^{-1}$ ) a 1 g NaCl. Následuje třístupňová extrakce do 4 ml ethylacetátu, při které se po 10 min třepání odebírá horní vrstva a přenesení do čisté zkumavky, kde se extrakty od daného vzorku spojí a vysuší se bezvodým síranem sodným (1 g, 10 min). Pak se roztok dekantuje do čisté zkumavky a odfouká pod dusíkem při 30 °C. Odparek se ze zkumavky vymývá ve třech krocích: 2×400  $\mu\text{l}$  methanolu a 300  $\mu\text{l}$  acetonu, a poté se ze zkumavky přenesení do GC-MS vialky, vysuší pod dusíkem při 30 °C a nakonec se rozpustí v 300  $\mu\text{l}$  silylační směsi (BSTFA:pyridin:TMCS, 25:25:1). Vialka se uzavírá pod dusíkem a 20 min udržuje při 70 °C.

### 11.3.2 Plasma

Vzorek plasmy (SRM 1950 [161]) je připraven podle McGaw a kol. [117]. 100  $\mu\text{l}$  vzorku je smícháno s interními standardy: 10  $\mu\text{l}$  MMAD (1,04  $\text{mmol.l}^{-1}$ ) a 10  $\mu\text{l}$  CLAC (1  $\text{mmol.l}^{-1}$ ). Po 15 min při 4 °C je přidáno 350  $\mu\text{l}$  roztoku methanolu (3:1, methanol:voda) a PHED (0,2  $\text{mmol.l}^{-1}$  v methanolu) a dobře promícháno, po 30 min při -20 °C. Vzorky jsou po 10 min centrifugovány při 4 °C (14 000 g). 350  $\mu\text{l}$  supernatantu je odpipetováno do čisté vialky a vysušeno pod proudem dusíku při 30 °C. Odparek je derivatizován ve dvou krocích: (1) 10  $\mu\text{l}$  EtOx (56 mg/ml v pyridinu) a 50  $\mu\text{l}$  pyridinu, 1 h při 50 °C; (2) 50  $\mu\text{l}$  MSTFA (45 min při 50 °C).

## 11.4 Metody

Stejný, čerstvě připravený, vzorek moči a plasmy je analyzován prostřednictvím Pegasus 4D GC×GC-TOF MS (LECO) pomocí různých kolonových uspořádání. Kolony první a druhé dimenze (všechny od SGE, s výjimkou SLB-IL59 od Sigma-Aldrich) jsou spojeny pomocí SiTite konektoru (SGE), modulace se vždy provádí na méně polární koloně, aby se předešlo ulpívání polárních látek v modulátoru. Rozměr kolony první dimenze je vždy stejný: délka 30 m, vnitřní průměr 0,25 mm, tloušťka filmu stacionární fáze 0,25  $\mu\text{m}$ . Parametry kolony druhé dimenze jsou pro moč a plasmu shrnuty v Tab. 10 a 11.

Analytický systém je vybaven termálním modulátorem s 2 horkými a 2 studenými tryskami. Studené trysky jsou chlazeny pomocí CFM (z angl. „consumable-free“) systému. 1  $\mu\text{l}$  vzorku je v režimu bez děliče toku (angl. „splitless“) je při 40 °C nastříknut do PTV injektoru (Gerstel CIS4), po 20 s je teplota rychlostí 10°C.min<sup>-1</sup> zvýšena na 250 °C, kde je 90 s udržována konstantní. Pak je rychlostí 10 °C.min<sup>-1</sup> teplota zvýšena na 350 °C, po 3 min při této teplotě se injektor vrací na výchozích 40°C.

Jako nosný plyn slouží helium v režimu konstantního průtoku (1 ml.min<sup>-1</sup>). Teplotní program termostatu primární kolony začíná na 80 °C, pak je teplota rychlostí 4 °C.min<sup>-1</sup> zvýšen na 130 °C, následně rychlostí 6 °C.min<sup>-1</sup> na 200 °C a nakonec rychlostí 12 °C.min<sup>-1</sup> na 285 °C, kde je 10 min udržována konstantní. Teplota modulátoru je nastavena vždy 35 °C nad aktuální teplotu primárního termostatu, modulační perioda na 4 s. Délka horkého pulsu se po 1894 s mění z 0,5 s na 0,75 s. Teplota termostatu sekundární kolony je o 5 °C vyšší než primární.

Do hmotnostního spektrometru jsou analyty převáděny při 250°C. Při stejné teplotě pracuje i iontový zdroj. Jedná se o elektronovou ionizaci s urychlujícím potenciálem 70 eV. Data se sbírají po uplynutí úvodních 250 s analýzy, a to v rozsahu 35-650  $m/z$  s frekvencí 100 Hz. Napětí na detektoru činí -1750 V. Analýza dat se provádí pomocí ChromaTOF software v. 4.24 (LECO), MS Excel, R-software (v. 3.3.0), and MATLAB (v. R2018a).

## 11.5 Výsledky a diskuse

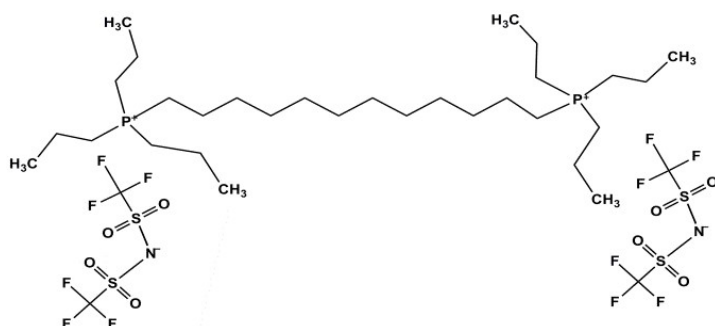
### 11.5.1 Určení ortogonálního kolonového uspořádání pro GC×GC analýzu moči

Většina GC×GC analýz moči připravené pomocí trimethylsilylace je prováděna pomocí normálního kolonového uspořádání, jmenovitě pomocí setu BPX5/BPX50 nebo jejich komerčních analog (Tab. 9). Nepolární kolony typu BPX5 se v první dimenzi používají díky dlouholetým zkušenostem z 1D GC, polární kolony druhé dimenze nabízejí jinou selektivitu. Zatímco retence analytů v první dimenzi odpovídá na základě Londonových disperzních sil jejich těkavosti, v druhé dimenzi se řídí interakcemi dipól-dipól. Kolonu první dimenze opouštějí látky stejné těkavosti, vstupují do kolony druhé dimenze, kde je separace velmi krátká, a díky tomu víceméně isothermální (podíl Londonových disperzních sil na separaci analytů v druhé dimenzi je tak zanedbatelný). Proto je takovéto spojení kolon považováno za ortogonální [162].

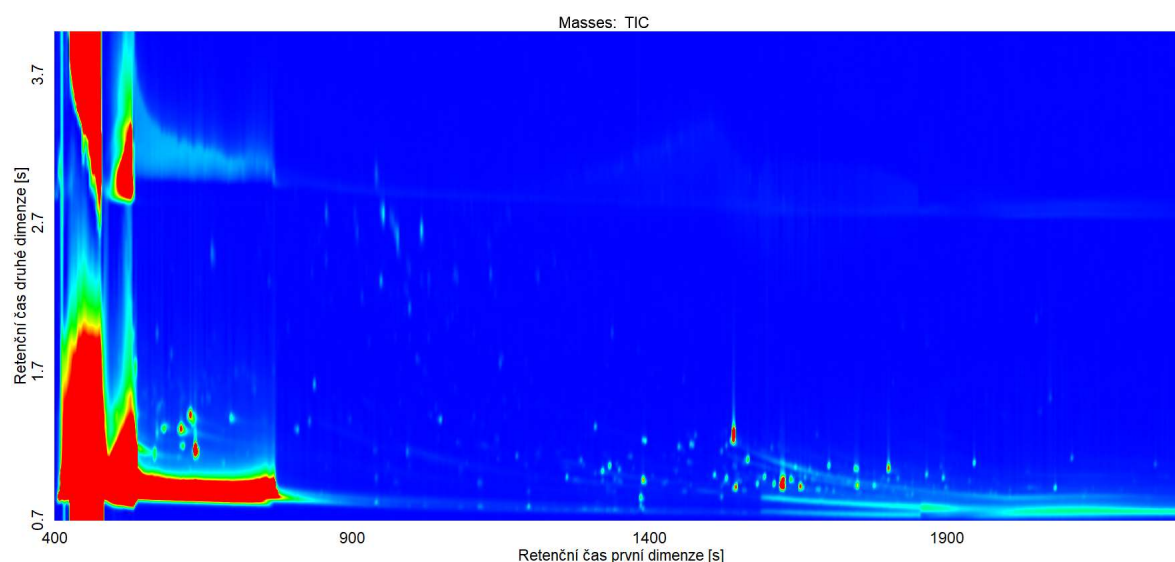
O použití jiných kolonových uspořádání k analýze metabolitů moči připravených trimethylsilylací není v literatuře žádná informace. Proto byl vzorek moči analyzován pomocí šesti různých kombinací kolon: dvou kombinací kolony nepolární s polární (BPX5/SolGel-WAX, BPX5/SLB-IL59), dvou reverzních setů kolon polárních s nepolárními (BPX50/BPX5, SolGel-WAX/BPX-5),

jednoho uspořádání středně polární kolony s polární kolonou (BPX50/SolGel-WAX) a jedné kombinace polární kolony se středně polární (SLB-IL59/BPX50).

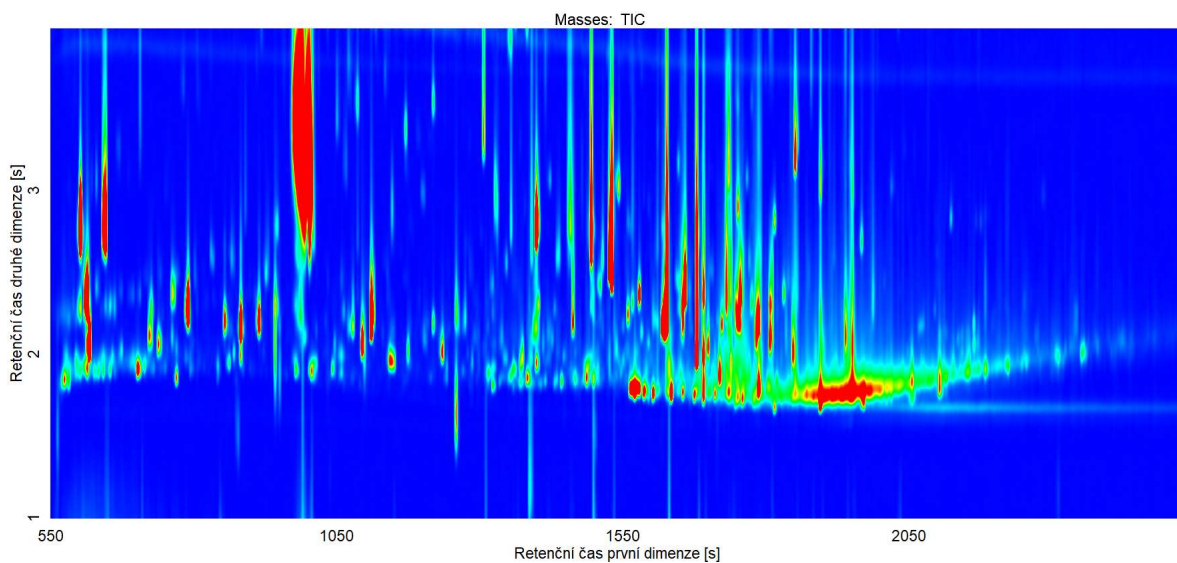
Kombinace polární kolony s nepolární ani polární se středně polární kolonou neposkytuje účelnou separaci. U kolony SLB-IL59, patřící mezi kolony s iontovou kapalinou (Obr. 45), je to způsobeno extrémní retencí analytů, zapříčiněnou hned třemi různými separačními mechanismy: interakcí dipól-indukovaný dipól, dipól-dipól a vodíkovými vazbami [163]. Když je takto kolona instalována v první dimenzi, nedochází k téměř žádné separaci (Obr. 46). V druhé dimenzi vyvolává extrémní chvostování píků (Obr. 47), které znemožňuje identifikaci řady analytů (Obr. 48). Chvostování je zapříčiněno nižší deaktivací křemičitanového povrchu kolony [164] a příliš tenkou vrstvou stacionární fáze působící nízkou zatžitelností kolony.



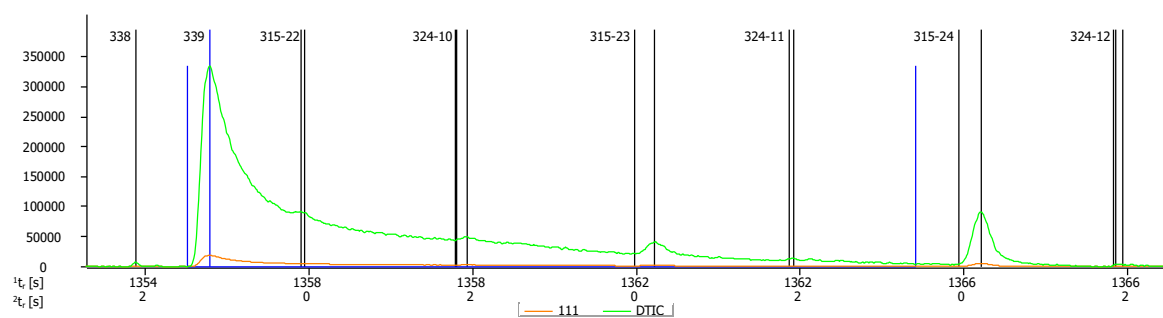
Obr. 45: Struktura stacionární fáze SLB IL-59 [165]



Obr. 46: GCxGC separace moči pomocí kombinace kolon SLB-IL59/BPX50 (Set A, Tab. 10)

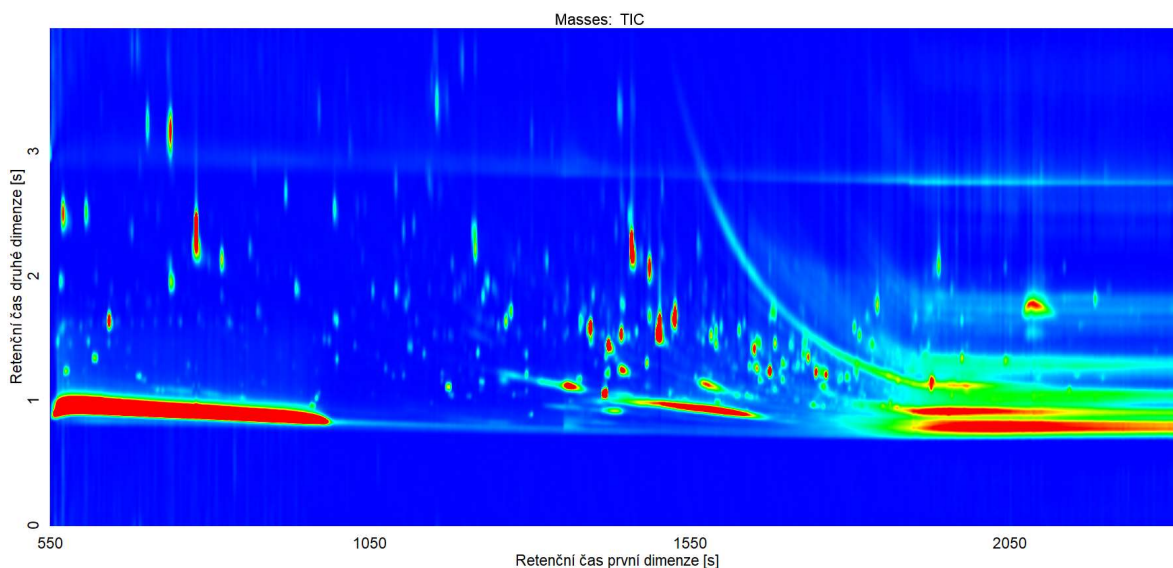


Obr. 47: GCxGC separace moči pomocí kombinace kolon BPX5/SLB-IL59 (Set B, Tab. 10)



Obr. 48: Tvar píku adipátu-2TMS (pík č. 339) analyzovaného pomocí setu B (Tab. 10). Černé přímky kolmé k ose x značí vrcholy píků, modré vymezují píku adipátu. Je zobrazen specifický fragment adipátu ( $m/z$  111) a dTIC (TIC, který po dekonvoluci náleží píku č. 339).





Obr. 49: GCxGC separace moči pomocí kombinace kolon SolGel-WAX/BPX5 (Set M, Tab. 10)

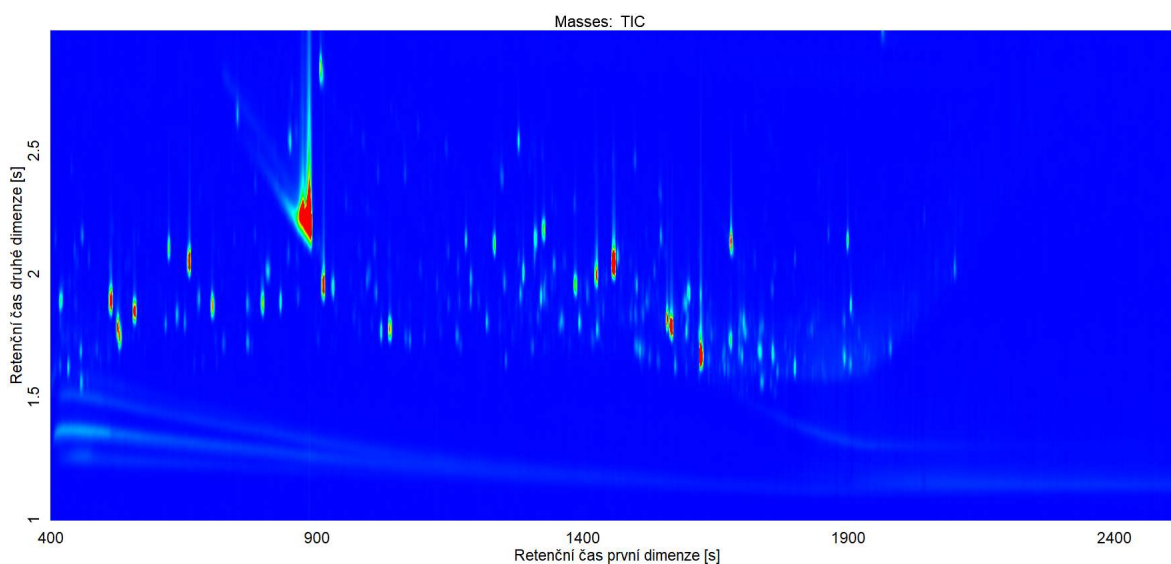
Kolona SolGel-WAX, polární kolona na bázi polyethylenglykolu, také vykazuje vysokou retenci analytů (Obr. 49) díky přítomnosti dvou volných elektronových párů kyslíku, které se podílejí na tvorbě vodíkových vazeb a dipólových interakcí [142]. Díky přítomnosti nepolárních řetězců polysiloxanu (SolGel), který je výhodný pro vyšší stabilitu kolony, je schopna vázat analyty také prostřednictvím Londonových dispersních sil. Přestože je kolona SolGel-WAX o něco účinnější než kolona SLB-IL59, je zřejmé, že polární kolony nejsou vhodné pro separaci látek připravených jako trimethylsilylderiváty. Analyty mají díky nekompletní derivatizaci analytů vysokou retenci. Pokud analyt obsahuje více kyselých vodíků, je v některých případech nahrazen trimethylsilylovou skupinou pouze některý z nich. Tento volný vodíkový atom pak silně interaguje se stacionární fází a vede k vysoké retenci analytů. Ze stejného důvodu tyto polární kolony také vysoce zadržují rozpouštědla (Obr. 46 a 49).

U zbývajících kolonových uspořádání byla vypočítána ortogonalita pomocí ASCA (Tab. 10). Všechny sety obsahující v druhé dimenzi kolonu s 0,25  $\mu\text{m}$  stacionární fází vykazují, ve srovnání se stejným typem stacionární fáze v tenčí vrstvě vyšší ortogonalitu. Z tohoto jevu lze usuzovat, že silnější vrstva stacionární fáze má výrazný vliv na retenci analytů. Jako nejvíce ortogonální je hodnoceno kolonové uspořádání BPX50/SolGel-WAX (set J,  $\alpha = 68,5\%$ ). Chromatogramy jsou uvedeny na Obr. 50 – 56.

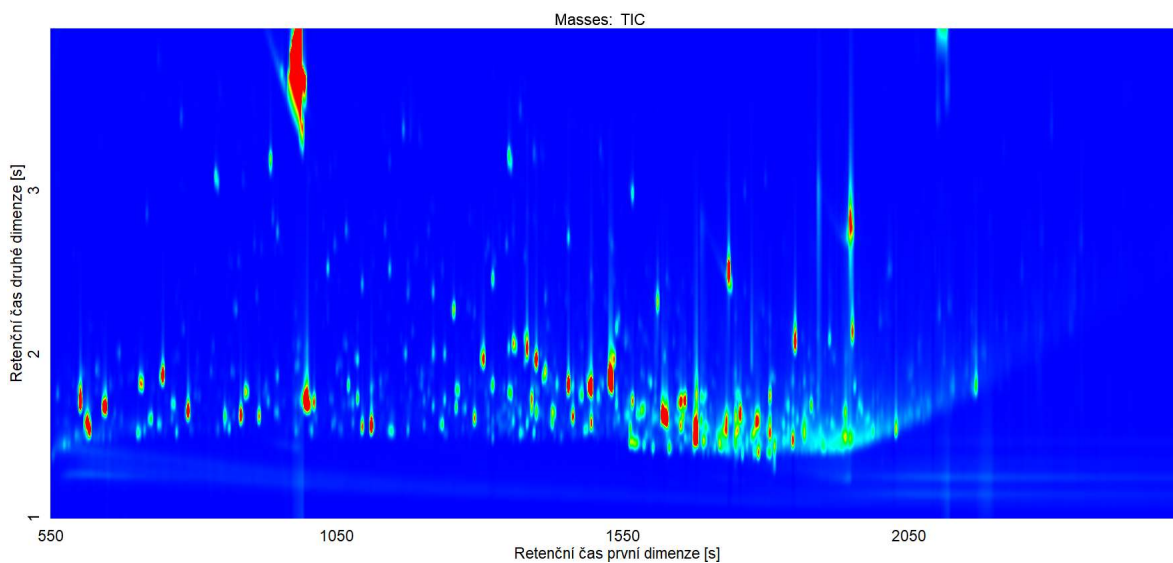
Tab. 10: Kolonová uspořádání použitá pro analýzu moči a hodnota ortogonalitě spočítaná pomocí ASCA

Set	Kolonová uspořádání ( <sup>1</sup> D/ <sup>2</sup> D)	Parametry kolony druhé dimenze (délka [m] – průměr [mm] – film [μm])	O [%]
A	SLB-IL59/BPX50	1,3 – 0,25 – 0,20	-
B	BPX5/SLB-IL59	2,0 – 0,10 – 0,08	-
C	BPX5/SolGel-WAX	1,2 – 0,10 – 0,10	56,0
H	BPX50/BPX5*	1,2 – 0,10 – 0,10	45,1
I	BPX50/SolGel-WAX	1,6 – 0,10 – 0,10	44,3
J	BPX50/SolGel-WAX	1,2 – 0,25 – 0,25	68,5
K	BPX5/SolGel-WAX	1,2 – 0,25 – 0,25	65,9
L	BPX50/BPX5*	1,3 – 0,25 – 0,25	64,3
M	SolGel-WAX/BPX-5*	1,3 – 0,10 – 0,10	-
N	BPX5/BPX50	1,5 – 0,10 – 0,10	56,9

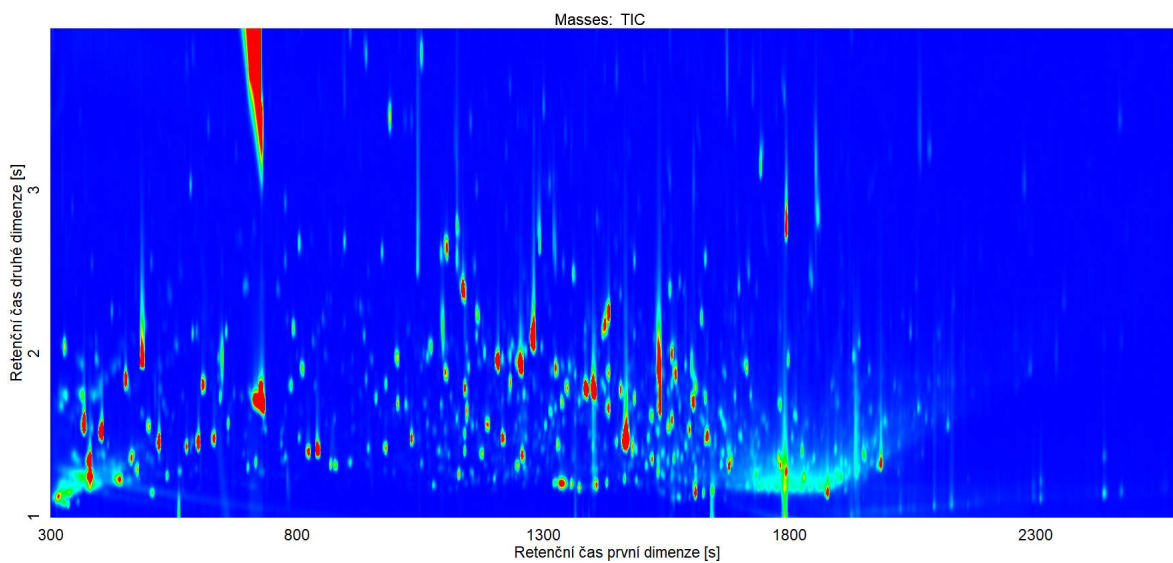
\* Modulace na koloně druhé dimenze



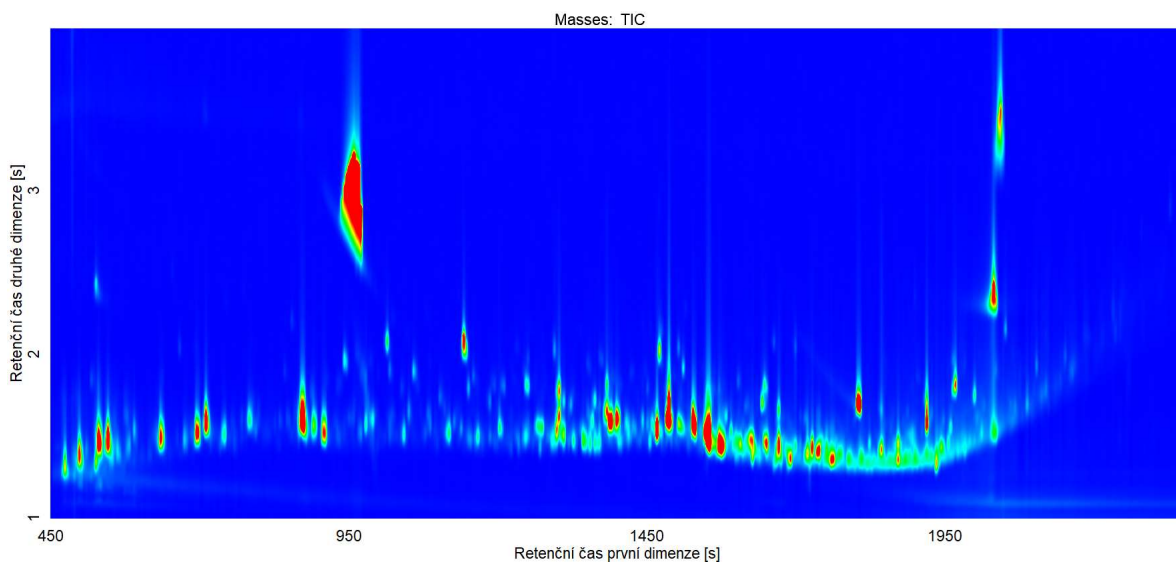
Obr. 50: GCxGC separace moči pomocí kombinace kolon BPX5/BPX50 (Set N, Tab. 10)



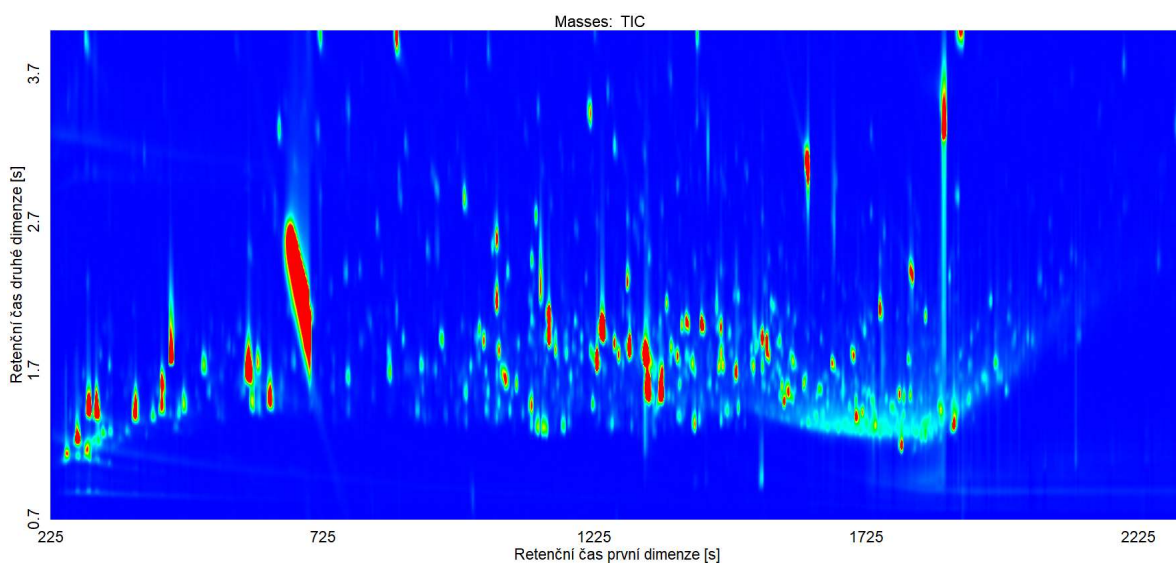
Obr. 51: GC×GC separace moči pomocí kombinace kolon BPX5/SolGel-WAX (Set C, Tab. 10, průměr kolony druhé dimenze 0,10 mm)



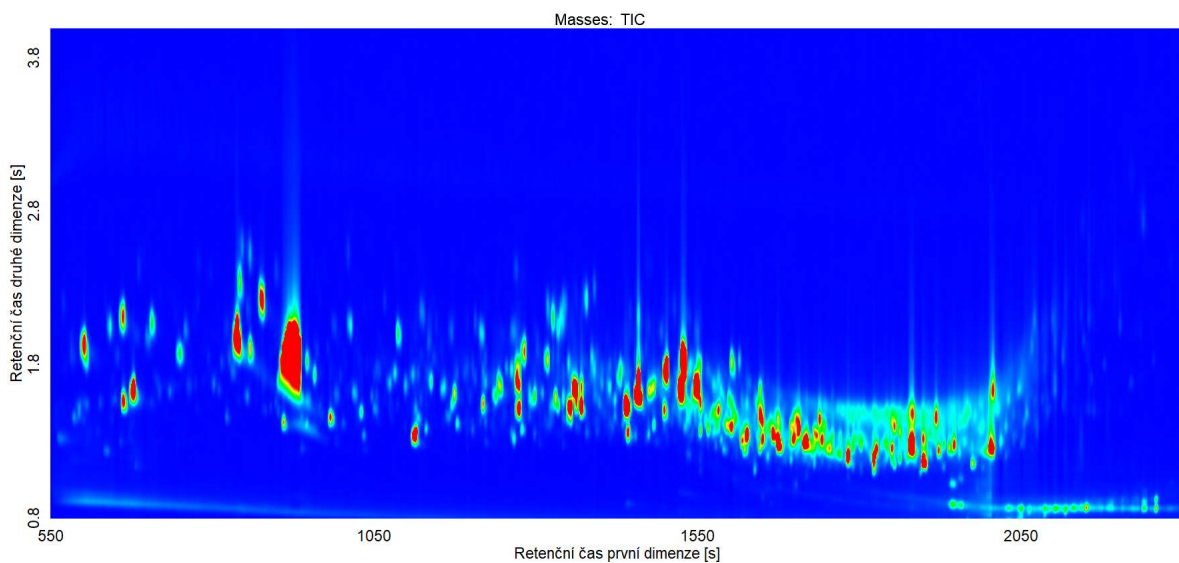
Obr. 52: GC×GC separace moči pomocí kombinace kolon BPX5/SolGel-WAX (Set K, Tab. 10, průměr kolony druhé dimenze 0,25 mm)



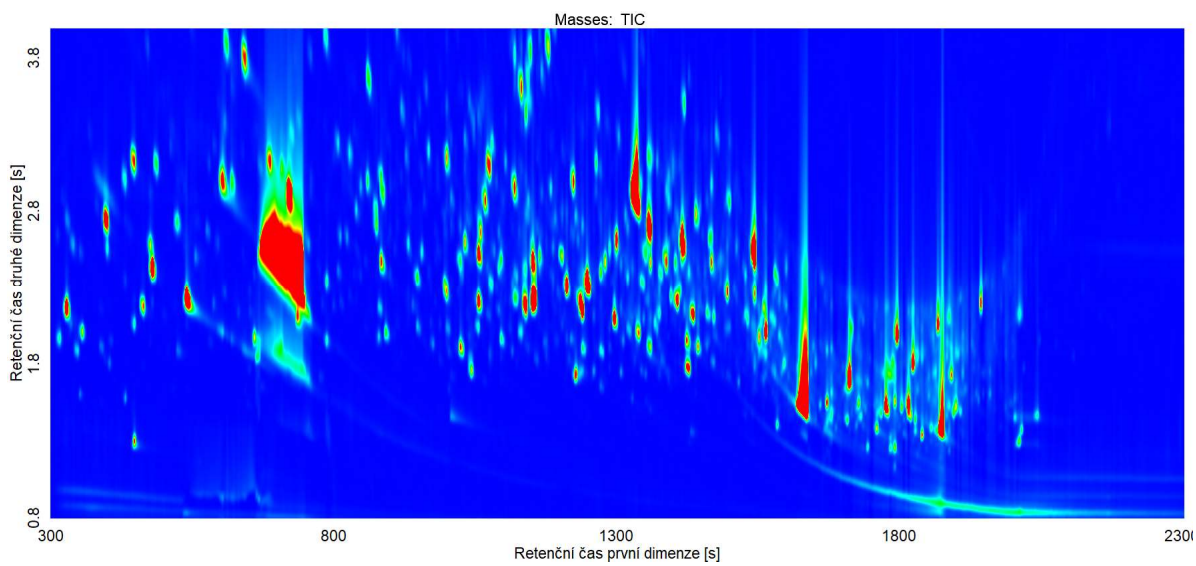
Obr. 53: GC×GC separace moči pomocí kombinace kolon BPX50/SolGel-WAX (Set I, Tab. 10, průměr kolony druhé dimenze 0,10 mm)



Obr. 54: GC×GC separace moči pomocí kombinace kolon BPX50/SolGel-WAX (Set J, Tab. 10, průměr kolony druhé dimenze 0,25 mm)



Obr. 55: GC×GC separace moči pomocí kombinace kolon BPX50/BPX5 (Set H, Tab. 10, průměr kolony druhé dimenze 0,10 mm)



Obr. 56: GC×GC separace moči pomocí kombinace kolon BPX50/BPX5 (Set L, Tab. 10, průměr kolony druhé dimenze 0,25 mm)

### 11.5.2 Určení ortogonálního kolonového uspořádání pro GC×GC analýzu plasmy

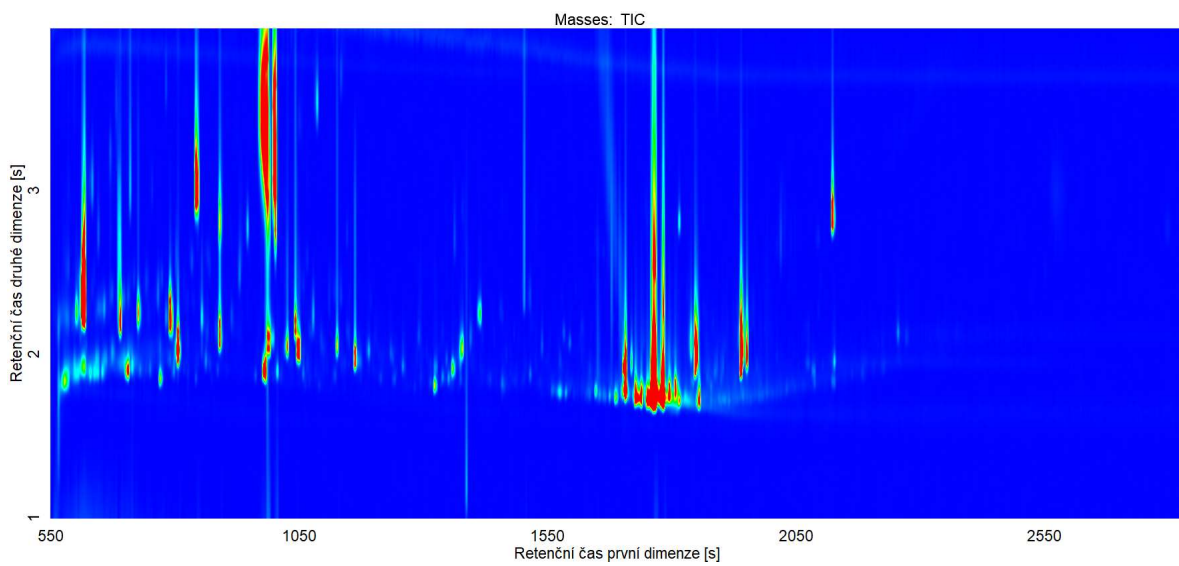
GC×GC analýza plasmy je kromě tradičního kolonového uspořádání odpovídajícímu chemismu stacionárních fází BPX5/BPX50 prováděna také pomocí reverzního uspořádání BPX50/BPX5, anebo pomocí RTX200, která obsahuje středně polární trifluoropropylmethylpolysiloxan. Trifluoropropylmethylpolysiloxan interaguje s volnými elektronovými páry a  $\pi$ -elektronovými systémy, proto kolona zvýšeně zadržuje nenasycené sloučeniny, aromáty, aldehydy, ketony, halogenderiváty a nitrosločeniny [142]. Použití iontových kapalin ani polyethylenglykolových kolon pro analýzu trimethylsilylovaných metabolitů moči literaturou doloženo není, proto jim v této práci byla věnována pozornost.

Stejně jako je tomu u analýzy moči, kolony druhé dimenze se silnější vrstvou stacionární fáze vykazují v porovnání se tenčí vrstvou vyšší ortogonalitu (Tab. 11). Jako nejlepší, co se ortogonality týče, je hodnocen set BPX50/BPX5 (Obr. 63). Tento závěr je v souladu s prací Koek a kol. [140]. Polyethylenglykolová kolona ani kolona na bázi iontové kapaliny není pro separaci v normálním uspořádání přínosem. SLB-IL59 u kolony druhé dimenze také u plasem způsobuje chvostování píků (Obr. 57). Při použití kolony SolGel-WAX ve druhé dimenzi se na chromatogramu objevuje méně píků s vyšší retencí na této koloně. Je tomu tak nejspíš proto, že plasma je ve srovnání s močí celkově jednodušší materiál (obsahuje méně metabolitů detekovatelných pomocí GC-MS a v nižší koncentraci), a tak k nekompletní derivatizaci dochází zřídka. Analytů obsahujících kyselou vodíky, které by mohly interagovat s polyethylenglykolovou stacionární fází, je proto velmi málo. Reverzní uspořádání, kdy by SolGel-WAX a SLB-IL59 byly kolonami první dimenze, nejsou na základě výsledků analýzy moči v případě plasem testovány.

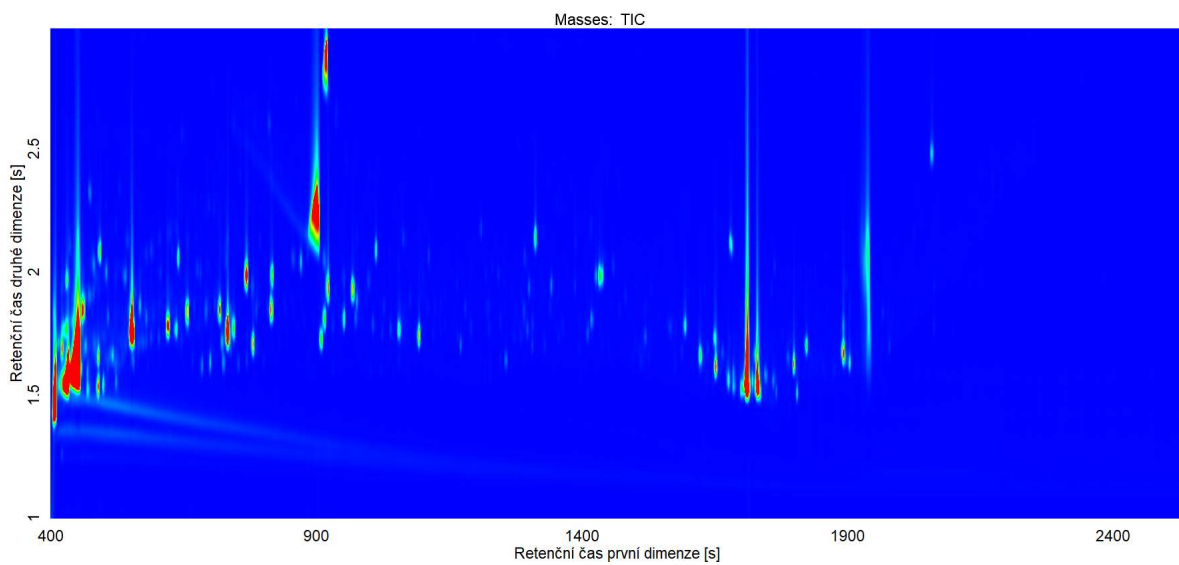
Tab. 11: Kolonová uspořádání použitá pro analýzu plasmy

Set	Kolonová uspořádání ( <sup>1</sup> D/ <sup>2</sup> D)	Parametry kolony druhé dimenze (délka [m] – průměr [mm] – film [μm])	O [%]
A	BPX5/BPX50	1,5 – 0,10 – 0,10	51,0
B	BPX5/SLB-IL59	2,0 – 0,10 – 0,08	53,5
C	BPX50/SolGel-WAX	1,6 – 0,10 – 0,10	44,7
D	BPX50/SolGel-WAX	1,2 – 0,25 – 0,25	56,7
E	BPX5/SolGel-WAX	1,2 – 0,25 – 0,25	58,4
F	BPX50/BPX5*	1,3 – 0,25 – 0,25	62,4
G	BPX5/SolGel-WAX	1,2 – 0,10 – 0,10	40,6

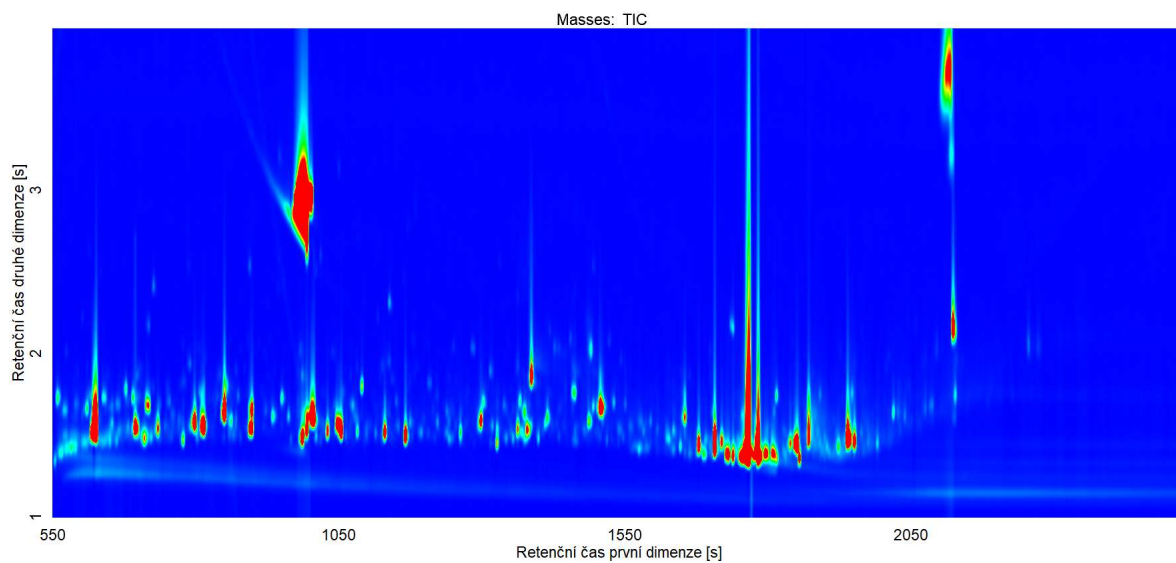
\* Modulace na koloně druhé dimenze



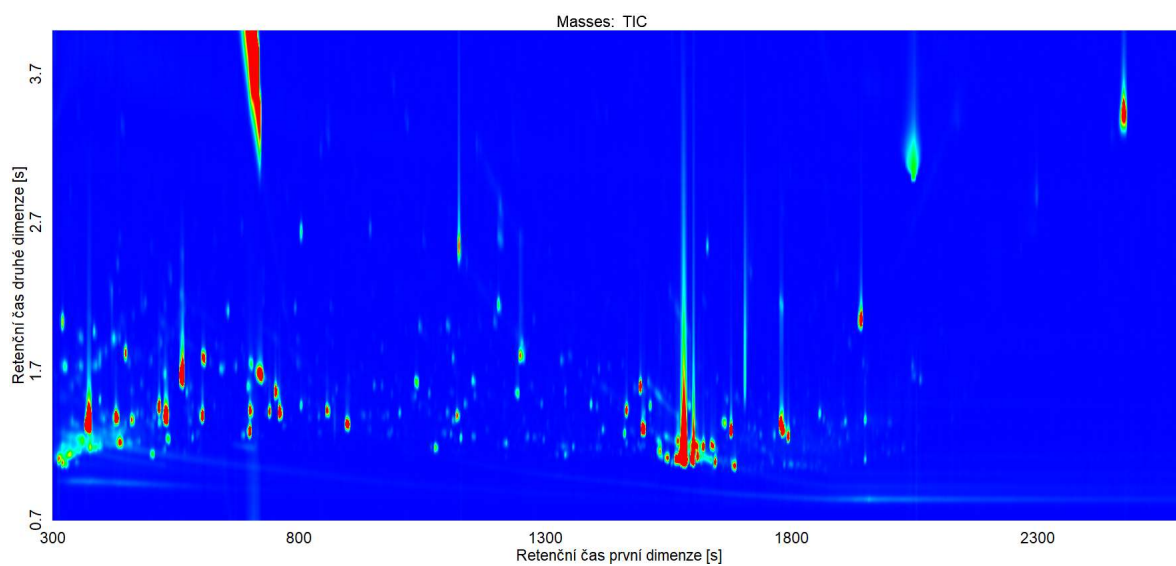
Obr. 57: GC×GC separace plasmy pomocí kombinace kolon BPX5/SLB-IL59 (Set B, Tab. 11)



Obr. 58: GC×GC separace plasmy pomocí kombinace kolon BPX5/BPX50 (Set A, Tab. 11)

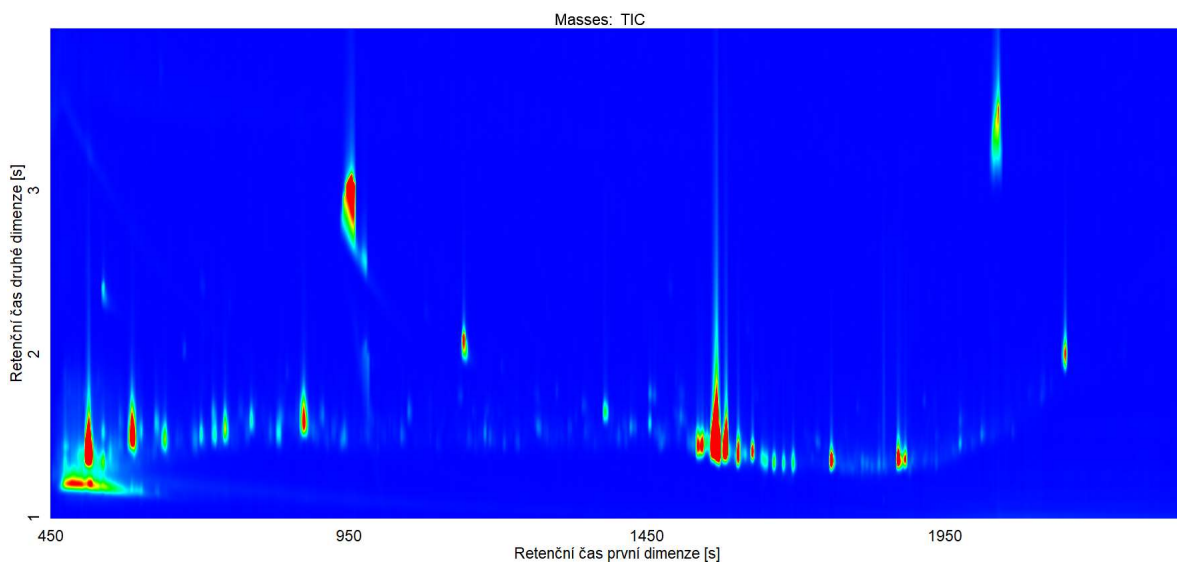


Obr. 59: GC×GC separace plasmu pomocí kombinace kolon BPX5/SolGel-WAX (Set G, Tab. 11, průměr kolony druhé dimenze 0,10 mm)

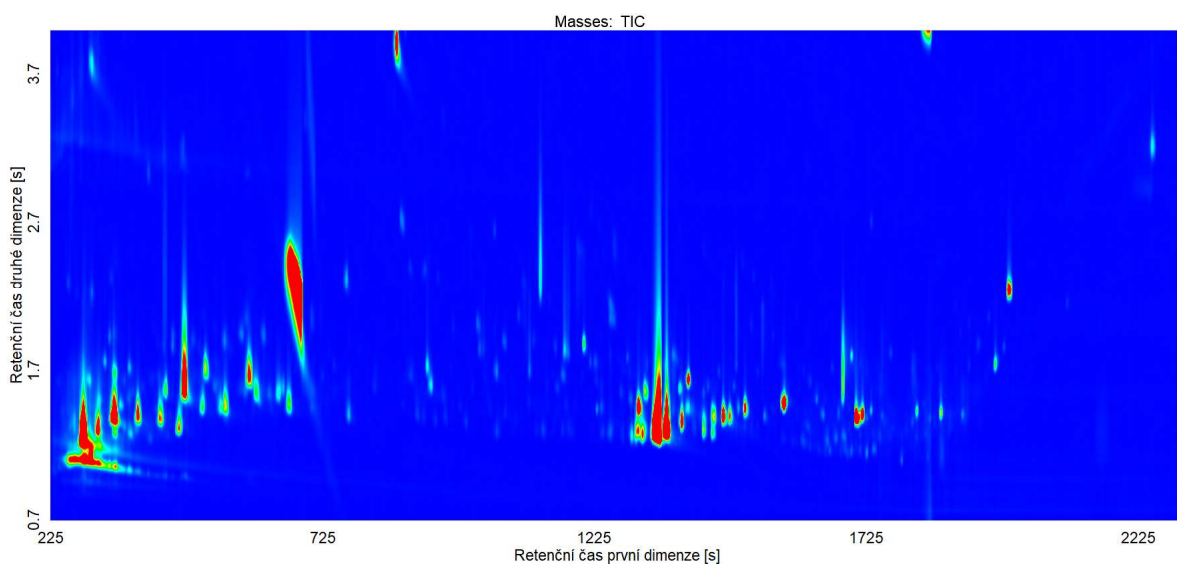


Obr. 60: GC×GC separace plasmu pomocí kombinace kolon BPX5/SolGel-WAX (Set E, Tab. 11, průměr kolony druhé dimenze 0,25 mm)

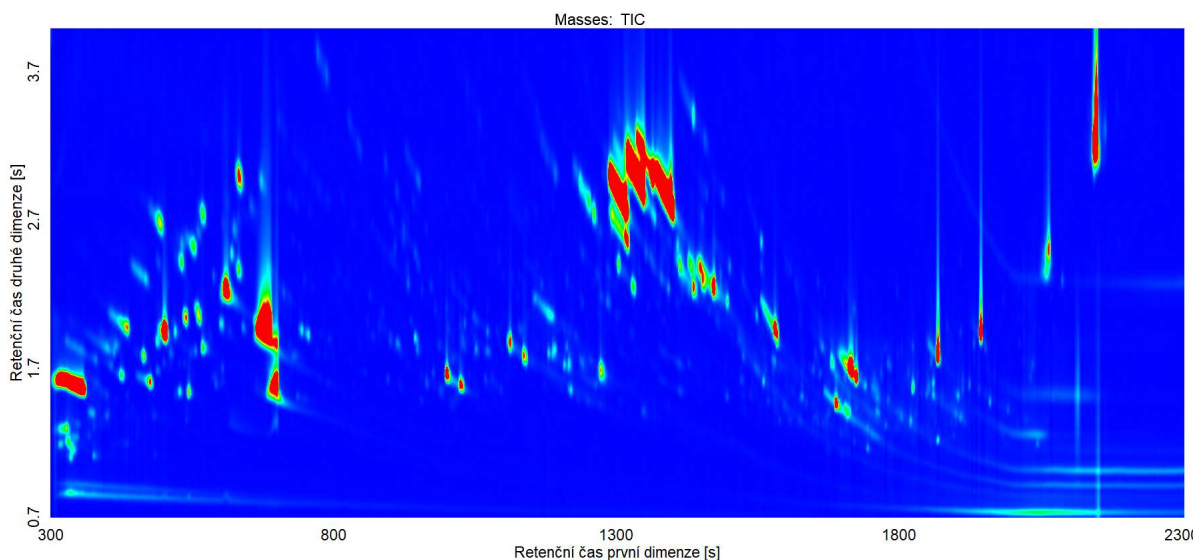




Obr. 61: GC×GC separace plasmy pomocí kombinace kolon BPX50/SolGel-WAX (Set C, Tab. 11, průměr kolony druhé dimenze 0,10 mm)



Obr. 62: GC×GC separace plasmy pomocí kombinace kolon BPX50/SolGel-WAX (Set D, Tab. 11, průměr kolony druhé dimenze 0,25 mm)



Obr. 63: GC×GC separace plasmy pomocí kombinace kolon BPX50/BPX5 (Set F, Tab. 11, průměr kolony druhé dimenze 0,25 mm)

### 11.5.3 Zamyšlení nad ortogonalitou v optimalizaci GC×GC systémů

Ortogonalita je jedním z parametrů používaných k optimalizaci dvojdimenzionálních chromatografických separačních systémů. Zůstává otázkou, zda je parametrem dostatečným. Například pro separaci moči bylo nejvíce ortogonální kombinací kolon shledáno uspořádání J (Tab. 10). Z Obr. 50-56 je patrné, že separace stejného vzorku moči pomocí různých kolonových uspořádání se liší i dalšími parametry. V Tab. 12 je uvedena šířka píku v polovině výšky a odstup signálu od šumu (S/N) pro adipát-2TMS a počet nalezených píků potenciálních metabolitů. Pokud se prostřednictvím geometrického průměru zkombinují do výsledného deskriptoru nazývaného *OS* (optimalizátor systému), je jako nejlepší GC×GC systém pro separaci moči vyhodnocen systém obsahující kolonové uspořádání K.

Druh a počet látek použitých pro stanovení S/N a šířky píku v polovině výšky je dalším kritériem, které je třeba zvážit. Ve výše uvedeném modelovém příkladu byl zvolen adipát, protože se nachází přibližně ve středu analytického okna. Navíc je to i diagnosticky významná látka. V moči však zastupuje pouze jednu skupinu přítomných analytů. Pro správnou charakterizaci separace by bylo třeba vybrat z každé skupiny minimálně jeden analyt. Navíc je zde i otázka vlivu koeluce vybrané charakteristické látky z dané skupiny s jinými analyty či složkami matrice, která má vliv na šířku píku a S/N. Tento

fenomén je však v deskriptoru *OS* kompenzován započítáním počtu nalezených potenciálních metabolitů.

Tab. 12: Chromatografické parametry separací moči pomocí různých kolonových uspořádání: *O* - ortogonalita [%],  $h_{1/2}$  - šířka píku adipátu (2 TMS) [s], *S/N* – odstup signálu od šumu, *P* – počet píků v separaci

Set	O [%]	$h_{1/2}$ [s]	S/N	P	OS
C	56,0	0,0697	13585	1251	90,25
H	45,1	0,0906	6796	1020	72,95
I	44,3	0,0878	7366	788	68,93
J	68,5	0,0854	12763	1116	95,54
K	65,9	0,0659	15126	1442	98,65
L	64,3	0,0970	9864	1158	91,87
N	56,9	0,0610	2049	417	41,50

## 11.6 Závěr

Nejvyšší ortogonalitu pro GC×GC analýzu moči poskytuje set BPX50/SolGel-WAX, zatímco pro analýzu plasmu je jím set BPX50/BPX5. V obou případech znamená použití kolony druhé dimenze s větším průměrem a tlustší vrstvou stacionární fáze výrazný nárůst retence analytů a tím i nárůst ortogonalit systému.

To, že se nejvíce ortogonální uspořádání kolon pro plasmu a moč liší, je pravděpodobně způsobeno rozdílným charakterem obou biologických materiálů. Protože koncentrace analytů v plasmě je řízena homeostatickými mechanismy, rozdíl v koncentraci analytů není tak velký. Podíl nedostatečně derivatizovaných analytů kvůli nadměrné spotřebě derivatizačního činidla je velmi malý. Proto mají analyty méně kyselých vodíků, které by mohly interagovat s polethylenglykolovou stacionární fází. Hlavní odlišující vlastností je kromě velikosti molekuly přítomnost  $\pi$ -elektronů.

Oproti tomu moč obsahuje řadu látek, které se do ní jako nepotřebné zplodiny metabolismu z těla vylučují ve vysoké koncentraci. Příkladem je močovina, hydroxyhippuráty aj. Proto je při derivatizaci metabolitů v moči spotřeba derivatizačního činidla vyšší, stejně jako zastoupení neúplně derivatizovaných látek s kyselými vodíky. SolGel-WAX kolona nainstalovaná v druhé dimenzi je proto pro analýzu moči výhodou. Při současném použití kolony BPX50 jsou první dimenzi jsou analyty separovány na základě Londonových disperzních sil a  $\pi$ - $\pi$  interakcí (tedy zcela odlišně).

## 12 Závěr

V souladu se stanovenými cíli je v rámci předložené práce vyvinut nový přístup přípravy vzorku moči pro metabolickou analýzu, který využívá ureasy imobilizované na magnetické částice. Tento přístup ve srovnání s ostatními testovanými přístupy (volná ureasa, kyselá extrakce do ethylacetátu a vzorky bez úpravy) minimálně mění metabolom vzorku (co do počtu vnesených artefaktů a změněných hladin metabolitů). Poskytuje také nízkou analytickou variabilitu koncentrací (2,5 %) a dobrou reprodukovatelnost retenčních indexů (pod 0,01 %).

V rámci zaměření na vícerozměrné chromatografické separační techniky je vyvinut nový způsob škálování dvojrozměrného separačního prostoru, nazývaný WOSEL. Tento přístup vymezuje separační prostor na základě parametrů analytické metody: začátek je v obou dimenzích dán mrtvým časem, konec posledním okamžikem sběru dat. Při použití WOSEL je oproti tradičním způsobům škálování rozsah separačního prostoru, ve kterém je počítána ortogonalita, nezávislý na vzorku použitým k optimalizaci.

Rovněž je zaveden nový deskriptor ortogonality, tzv. ASCA, který umožňuje nalezení kombinace kolon s nejvyšší ortogonalitou pro separaci daného vzorku. ASCA je geometrickým průměrem nejlepšího nalezeného deskriptoru lokální ortogonality (SC) a globální ortogonality (AN). Umožňuje tedy optimalizaci maximálního využití separačního prostoru (globální ortogonality) a minimálního shlukování dat (lokální ortogonality) pomocí jediné proměnné.

Nakonec je pomocí ASCA vybrána nejvíce ortogonální kombinace kolon pro analýzu metabolomu moči. Je jí kolonové uspořádání BPX50/SolGel-WAX s průměrem kolony druhé dimenze 0,25 mm a tloušťkou vrstvy stacionární fáze 0,25  $\mu\text{m}$ . Oproti tomu pro analýzu plasmy je nejlepší set kolon BPX50/BPX5 také s průměrem kolony druhé dimenze 0,25 mm a tloušťkou vrstvy stacionární fáze 0,25  $\mu\text{m}$ .

## 13 Výhledy

Příprava moči na bázi imobilizované ureasy je slibným přístupem pro GC-MS metabolomiku. Vnáší do vzorku minimum artefaktů a je snadno proveditelná. Právě toto zjednodušení přípravy vzorku by mohlo vést až k plné automatizaci. Vlastní proces imobilizace na komerční magnetické částice je poněkud zdlouhavý (imobilizovaná ureasa je připravena k použití až 3. den) a zakoupení mikročástic vyžaduje dodatečné náklady. Dostatečná skladovací stabilita imobilizované ureasy při (-80 °C) naznačuje možnost její komerční výroby a dodávání na zakázku. Nevýhodou by však mohla být její konečná cena. Levnější alternativu by mohla poskytnout domácí příprava magnetického nosiče, např. na bázi chitosanu [166], který je levný a připravuje se snadno. Bude však třeba vyzkoušet, zda nosič nebude interagovat s metabolity vzorku a výsledná získaná imobilizovaná ureasa bude dostatečně aktivní a stabilní. Další možností je příprava zesíťované ureasy s obsahem magnetitu, kterou by bylo možné skladovat lyofilizovanou a před použitím rekonstituovat, jak je to pro enzymy běžné.

Ve vzorcích zdravých kontrol připravených pomocí IUT byly necíleným GC-MS metabolickým přístupem nalezeny metabolity, které jsou u pacientů s GAI nedetekovatelné. Tento fenomén je pozorován i u jiných metabolických poruch. Předmětem další práce bude tyto látky identifikovat.

Při porovnání deskriptorů ortogonalita je jen okrajově zvažován vliv počtu píků v separaci, porovnáním 100 a 2500 píků. Jednotlivé deskriptory se však mohou lišit v citlivosti v oblasti malého, středního či velkého počtu separovaných píků. V některých aplikacích je 2D separačních technik, např. v LC×LC, je počet analytů nízký. Bylo by proto vhodné deskriptory (včetně ASCA) otestovat.

V oblasti výběru optimálního kolonového uspořádání je zde hodnocena pouze ortogonalita. Ortogonalita však není jediným důležitým aspektem dvojrozměrných chromatografických technik. Počet důležitých aspektů GC×GC separace a způsob jejich zkombinování dosud nebyl v literatuře řešen. Jedinou výjimkou je práce Nowika a kol. [167], která bere u přístupu NN v potaz i chromatografické rozlišení.

Při optimalizaci kolon je v této práci používáno generální schéma, které bylo aplikováno na všechna kolonová uspořádání. Je však všeobecně známo, že i jiné podmínky analýzy než pouhé chemické složení stacionární fáze, mají také vliv na ortogonalitu. V této práci jsou tedy kolonová uspořádání porovnána jen zevrubně (zejména za účelem porovnání deskriptorů) a nachystána pro další, detailní optimalizaci.

## 14 Seznam použité literatury

- [1] C.H. Johnson, J. Ivanisevic, G. Siuzdak, Metabolomics: beyond biomarkers and towards mechanisms, *Nat. Rev. Mol. Cell Biol.* 17 (2016) 451. <http://dx.doi.org/10.1038/nrm.2016.25>.
- [2] D.S. Wishart, Emerging applications of metabolomics in drug discovery and precision medicine, *Nat. Rev. Drug Discov.* 15 (2016) 473. <http://dx.doi.org/10.1038/nrd.2016.32>.
- [3] M. Mamas, W.B. Dunn, L. Neyses, R. Goodacre, The role of metabolites and metabolomics in clinically applicable biomarkers of disease, *Arch. Toxicol.* 85 (2011) 5–17. doi:10.1007/s00204-010-0609-6.
- [4] T.A. Clayton, J.C. Lindon, O. Cloarec, H. Antti, C. Charuel, G. Hanton, J.-P. Provost, J.-L. Le Net, D. Baker, R.J. Walley, J.R. Everett, J.K. Nicholson, Pharmaco-metabonomic phenotyping and personalized drug treatment., *Nature.* 440 (2006) 1073–7. doi:10.1038/nature04648.
- [5] J.K. Nicholson, J.C. Lindon, Systems biology: Metabonomics, *Nature.* 455 (2008) 1054–1056. doi:10.1038/4551054a.
- [6] H. Janečková, P. Wojtowicz, K. Hron, D. Friedecký, T. Adam, Nacílená metabolická analýza suchých krevních skvrn pro diagnostiku dědičných metabolických poruch, 20 (2012) 167–169.
- [7] L.D. Roberts, A.L. Souza, R.E. Gerszten, C.B. Clish, Targeted metabolomics., *Curr. Protoc. Mol. Biol.* Chapter 30 (2012) Unit 30.2.1-24. doi:10.1002/0471142727.mb3002s98.
- [8] N. Vinayavekhin, A. Saghatelian, Untargeted Metabolomics, in: *Curr. Protoc. Mol. Biol.*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2010: pp. 30.2.1-30.2.24. doi:10.1002/0471142727.mb3001s90.
- [9] K.A. Kaiser, C.E. Merrywell, F. Fang, C.K. Larive, *Metabolic Profiling*, Springer New York, New York, NY, 2018. doi:10.1007/978-1-4939-7643-0.
- [10] A. Zhang, H. Sun, P. Wang, Y. Han, X. Wang, Modern analytical techniques in metabolomics analysis, *Analyst.* 137 (2012) 293–300. doi:10.1039/C1AN15605E.
- [11] M. Giera, *Clinical Metabolomics*, Springer New York, New York, NY, 2018. doi:10.1007/978-1-4939-7592-1.
- [12] D.S. Wishart, C. Knox, A.C. Guo, R. Eisner, N. Young, B. Gautam, D.D. Hau, N. Psychogios, E. Dong, S. Bouatra, R. Mandal, I. Sinelnikov, J. Xia, L. Jia, J.A. Cruz, E. Lim, C.A. Sobsey, S. Shrivastava, P. Huang, P. Liu, L. Fang, J. Peng, R. Fradette, D. Cheng, D. Tzur, M. Clements, A. Lewis, A. De Souza, A. Zuniga, M. Dawe, Y. Xiong, D. Clive, R. Greiner, A. Nazyrova, R. Shaykhtudinov, L. Li, H.J. Vogel, I. Forsythe, HMDB : a knowledgebase for the human metabolome, 37 (2009) 603–610. doi:10.1093/nar/gkn810.

- [13] S. Bouatra, F. Aziat, R. Mandal, A.C. Guo, M.R. Wilson, C. Knox, T.C. Bjorndahl, R. Krishnamurthy, F. Saleem, P. Liu, Z.T. Dame, J. Poelzer, J. Huynh, F.S. Yallou, N. Psychogios, E. Dong, R. Bogumil, C. Roehring, D.S. Wishart, The Human Urine Metabolome, *PLoS One*. 8 (2013). doi:10.1371/journal.pone.0073076.
- [14] J. Ahn, J. Kim, J. Hwang, J. Song, K. Kim, H.-S. Cha, Urinary Metabolomic Profiling to Identify Potential Biomarkers for the Diagnosis of Behcet's Disease by Gas Chromatography/Time-of-Flight–Mass Spectrometry, *Int. J. Mol. Sci.* 18 (2017) 2309. doi:10.3390/ijms18112309.
- [15] J. Xu, Y. Chen, R. Zhang, J. He, Y. Song, J. Wang, H. Wang, L. Wang, Q. Zhan, Z. Abliz, Global metabolomics reveals potential urinary biomarkers of esophageal squamous cell carcinoma for diagnosis and staging, *Sci. Rep.* 6 (2016) 1–10. doi:10.1038/srep35010.
- [16] A.W. Michell, D. Mosedale, D.J. Grainger, R.A. Barker, Metabolomic analysis of urine and serum in Parkinson's disease, *Metabolomics*. 4 (2008) 191–201. doi:10.1007/s11306-008-0111-9.
- [17] A. Zhang, H. Sun, P. Wang, Y. Han, X. Wang, Recent and potential developments of biofluid analyses in metabolomics, *J. Proteomics*. 75 (2012) 1079–1088. doi:10.1016/j.jprot.2011.10.027.
- [18] J. Jacyna, M. Kordalewska, M.J. Markuszewski, Design of Experiments in metabolomics-related studies: An overview, *J. Pharm. Biomed. Anal.* 164 (2019) 598–606. doi:10.1016/j.jpba.2018.11.027.
- [19] J. Schripsema, D. Dagnino, Metabolomics: Experimental Design, Methodology, and Data Analysis, in: *Encycl. Anal. Chem.*, Wiley, 2019: pp. 1–22. doi:10.1002/9780470027318.a9939.pub2.
- [20] K. Tanaka, A. West-Dull, D.G. Hine, T.B. Lynn, T. Lowe, Gas-chromatographic method of analysis for urinary organic acids. II. Description of the procedure, and its application to diagnosis of patients with organic acidurias., *Clin. Chem.* 26 (1980) 1847–53. <http://www.ncbi.nlm.nih.gov/pubmed/7438430>.
- [21] H. Janečková, K. Hron, P. Wojtowicz, E. Hlídková, A. Barešová, D. Friedecký, L. Žídková, P. Hornik, D. Behúlová, D. Procházková, H. Vinohradská, K. Pešková, P. Bruheim, V. Smolka, S. Šťastná, T. Adam, Targeted metabolomic analysis of plasma samples for the diagnosis of inherited metabolic disorders, *J. Chromatogr. A*. 1226 (2012) 11–17. doi:10.1016/j.chroma.2011.09.074.
- [22] L. Najdekr, D. Friedecký, R. Tautenhahn, T. Pluskal, J. Wang, Y. Huang, T. Adam, Influence of Mass Resolving Power in Orbital Ion-Trap Mass Spectrometry-Based Metabolomics, *Anal. Chem.* 88 (2016) 11429–11435. doi:10.1021/acs.analchem.6b02319.

- [23] S.G. Villas-Bôas, J. Højer-Pedersen, M. Åkesson, J. Smedsgaard, J. Nielsen, Global metabolite analysis of yeast: evaluation of sample preparation methods, *Yeast*. 22 (2005) 1155–1169. doi:10.1002/yea.1308.
- [24] L. Tang, P. Kebarle, Dependence of ion intensity in electrospray mass spectrometry on the concentration of the analytes in the electrosprayed solution, *Anal. Chem.* 65 (1993) 3654–3668. doi:10.1021/ac00072a020.
- [25] A.C. Schrimpe-Rutledge, S.G. Codreanu, S.D. Sherrod, J.A. McLean, Untargeted Metabolomics Strategies—Challenges and Emerging Directions, *J. Am. Soc. Mass Spectrom.* 27 (2016) 1897–1905. doi:10.1007/s13361-016-1469-y.
- [26] J. Lisec, F. Hoffmann, C. Schmitt, C. Jaeger, Extending the Dynamic Range in Metabolomics Experiments by Automatic Correction of Peaks Exceeding the Detection Limit, *Anal. Chem.* 88 (2016) 7487–7492. doi:10.1021/acs.analchem.6b02515.
- [27] W.B. Dunn, D. Broadhurst, P. Begley, E. Zelena, S. Francis-McIntyre, N. Anderson, M. Brown, J.D. Knowles, A. Halsall, J.N. Haselden, A.W. Nicholls, I.D. Wilson, D.B. Kell, R. Goodacre, Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry, *Nat. Protoc.* 6 (2011) 1060–1083. doi:10.1038/nprot.2011.335.
- [28] S. Castillo, I. Mattila, J. Miettinen, M. Orešič, T. Hyötyläinen, Data Analysis Tool for Comprehensive Two-Dimensional Gas Chromatography/Time-of-Flight Mass Spectrometry, *Anal. Chem.* 83 (2011) 3058–3067. doi:10.1021/ac103308x.
- [29] Z. Yu, H. Huang, A. Reim, P.D. Charles, A. Northage, D. Jackson, I. Parry, B.M. Kessler, Optimizing 2D gas chromatography mass spectrometry for robust tissue, serum and urine metabolite profiling, *Talanta*. 165 (2017) 685–691. doi:10.1016/j.talanta.2017.01.003.
- [30] J.C. Giddings, Two-dimensional separations: concept and promise., *Anal. Chem.* 56 (1984) 1258A-1260A, 1262A, 1264A passim. doi:10.1021/ac00276a003.
- [31] C.J. Venkatramani, J. Xu, J.B. Phillips, Separation orthogonality in temperature-programmed comprehensive two-dimensional gas chromatography., *Anal. Chem.* 68 (1996) 1486–92. doi:10.1021/ac951048b.
- [32] K. Banerjee, S.H. Patil, S. Dasgupta, D.P. Oulkar, S.B. Patil, R. Savant, P.G. Adsule, Optimization of separation and detection conditions for the multiresidue analysis of pesticides in grapes by comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry, *J. Chromatogr. A*. 1190 (2008) 350–357. doi:10.1016/j.chroma.2008.03.017.



- [33] H. Gu, Y. Huang, P.W. Carr, Peak capacity optimization in comprehensive two dimensional liquid chromatography: A practical approach, *J. Chromatogr. A.* 1218 (2011) 64–73. doi:10.1016/j.chroma.2010.10.096.
- [34] E. Van Gyseghem, I. Crosiers, S. Gourvénec, D. Massart, Y. Vander Heyden, Determining orthogonal and similar chromatographic systems from the injection of mixtures in liquid chromatography–diode array detection and the interpretation of correlation coefficients color maps, *J. Chromatogr. A.* 1026 (2004) 117–128. doi:10.1016/j.chroma.2003.10.116.
- [35] M.R. Pourhaghighi, M. Karzand, H.H. Girault, Orthogonality of Two-Dimensional Separations Based on Conditional Entropy, *Anal. Chem.* 83 (2011) 7676–7681. doi:10.1021/ac2017772.
- [36] G. Semard, V. Peulon-Agasse, A. Bruchet, J.-P.P. Bouillon, P. Cardinaël, Convex hull: A new method to determine the separation space used and to optimize operating conditions for comprehensive two-dimensional gas chromatography, *J. Chromatogr. A.* 1217 (2010) 5449–54. doi:10.1016/j.chroma.2010.06.048.
- [37] J. V. Seeley, S.K. Seeley, Multidimensional gas chromatography: Fundamental advances and new applications, *Anal. Chem.* 85 (2013) 557–578. doi:10.1021/ac303195u.
- [38] M. Gilar, P. Olivova, A.E. Daly, J.C. Gebler, Orthogonality of separation in two-dimensional liquid chromatography., *Anal. Chem.* 77 (2005) 6426–34. doi:10.1021/ac050923i.
- [39] N.E. Watson, J.M. Davis, R.E. Synovec, Observations on “orthogonality” in comprehensive two-dimensional separations., *Anal. Chem.* 79 (2007) 7924–7. doi:10.1021/ac0710578.
- [40] W. Nowik, S. Héron, M. Bonose, M. Nowik, A. Tchaplá, Assessment of two-dimensional separative systems using nearest-neighbor distances approach. Part 1: orthogonality aspects., *Anal. Chem.* 85 (2013) 9449–58. doi:10.1021/ac4012705.
- [41] M. Camenzuli, P.J. Schoenmakers, A new measure of orthogonality for multi-dimensional chromatography., *Anal. Chim. Acta.* 838 (2014) 93–101. doi:10.1016/j.aca.2014.05.048.
- [42] Z.-D. Zeng, H.M. Hugel, P.J. Marriott, A modeling approach for orthogonality of comprehensive two-dimensional separations., *Anal. Chem.* 85 (2013) 6356–63. doi:10.1021/ac400736v.
- [43] W. Steuer, I. Grant, F. Erni, Comparison of high-performance liquid chromatography, supercritical fluid chromatography and capillary zone electrophoresis in drug analysis., *J. Chromatogr.* 507 (1990) 125–40. doi:10.1016/S0021-9673(01)84188-1.
- [44] Z. Liu, D.G. Patterson, M.L. Lee, Geometric Approach to Factor Analysis for the Estimation of Orthogonality and Practical Peak Capacity in Comprehensive Two-Dimensional Separations, *Anal. Chem.* 67 (1995) 3840–3845. doi:10.1021/ac00117a004.

- [45] J.C. Giddings, Sample dimensionality: a predictor of order-disorder in component peak distribution in multidimensional separation., *J. Chromatogr. A.* 703 (1995) 3–15. doi:10.1016/0021-9673(95)00249-M.
- [46] S. Zhu, X. Lu, K. Ji, K. Guo, Y. Li, C. Wu, G. Xu, Characterization of flavor compounds in Chinese liquor Moutai by comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry, *Anal. Chim. Acta.* 597 (2007) 340–348. doi:10.1016/j.aca.2007.07.007.
- [47] B.J. Blaise, G. Correia, A. Tin, J.H. Young, A.-C. Vergnaud, M. Lewis, J.T.M. Pearce, P. Elliott, J.K. Nicholson, E. Holmes, T.M.D. Ebbels, Power Analysis and Sample Size Determination in Metabolic Phenotyping, *Anal. Chem.* 88 (2016) 5179–5188. doi:10.1021/acs.analchem.6b00188.
- [48] J. Villanueva, Differential exoprotease activities confer tumor-specific serum peptidome patterns, *J. Clin. Invest.* 116 (2005) 271–284. doi:10.1172/JCI26022.
- [49] E.P. Diamandis, Cancer Biomarkers: Can We Turn Recent Failures into Success?, *JNCI J. Natl. Cancer Inst.* 102 (2010) 1462–1467. doi:10.1093/jnci/djq306.
- [50] M. Rotter, S. Brandmaier, C. Prehn, J. Adam, S. Rabstein, K. Gawrych, T. Brüning, T. Illig, H. Lickert, J. Adamski, R. Wang-Sattler, Stability of targeted metabolite profiles of urine samples under different storage conditions, *Metabolomics.* 13 (2017) 4. doi:10.1007/s11306-016-1137-z.
- [51] B. Kamlage, S. Neuber, B. Bethan, S.G. Maldonado, A. Wagner-Golbs, E. Peter, O. Schmitz, P. Schatz, Impact of prolonged blood incubation and extended serum storage at room temperature on the human serum metabolome, *Metabolites.* 8 (2018) 6. doi:10.3390/metabo8010006.
- [52] Y. Yang, G. Dong, Z. Wang, J. Wang, Z. Zhang, J. Liu, Rumen and plasma metabolomics profiling by UHPLC-QTOF/MS revealed metabolic alterations associated with a high-corn diet in beef steers, *PLoS One.* 13 (2018) e0208031. doi:10.1371/journal.pone.0208031.
- [53] R. Wawrzyniak, A. Kosnowska, S. Macioszek, R. Bartoszewski, M. Jan Markuszewski, New plasma preparation approach to enrich metabolome coverage in untargeted metabolomics: plasma protein bound hydrophobic metabolite release with proteinase K, *Sci. Rep.* 8 (2018) 9541. doi:10.1038/s41598-018-27983-0.
- [54] B.F. Darst, R.L. Koscik, K.J. Hogan, S.C. Johnson, C.D. Engelman, Longitudinal plasma metabolomics of aging and sex, *Aging (Albany. NY).* 11 (2019) 1262–1282. doi:10.18632/aging.101837.

- [55] K.J. Li, N. Jenkins, G. Luckasen, S. Rao, E.P. Ryan, Plasma metabolomics of children with aberrant serum lipids and inadequate micronutrient intake, *PLoS One*. 13 (2018) e0205899. doi:10.1371/journal.pone.0205899.
- [56] N.P. Paynter, R. Balasubramanian, F. Giulianini, D.D. Wang, L.F. Tinker, S. Gopal, A.A. Deik, K. Bullock, K.A. Pierce, J. Scott, M.A. Martínez-González, R. Estruch, J.E. Manson, N.R. Cook, C.M. Albert, C.B. Clish, K.M. Rexrode, Metabolic Predictors of Incident Coronary Heart Disease in Women, *Circulation*. 137 (2018) 841–853. doi:10.1161/CIRCULATIONAHA.117.029468.
- [57] S.K.P. Lau, K.-C. Lee, S.O.T. Curreem, W.-N. Chow, K.K.W. To, I.F.N. Hung, D.T.Y. Ho, S. Sridhar, I.W.S. Li, V.S.Y. Ding, E.W.F. Koo, C.-F. Wong, S. Tam, C.-W. Lam, K.-Y. Yuen, P.C.Y. Woo, Metabolomic Profiling of Plasma from Patients with Tuberculosis by Use of Untargeted Mass Spectrometry Reveals Novel Biomarkers for Diagnosis, *J. Clin. Microbiol.* 53 (2015) 3750–3759. doi:10.1128/JCM.01568-15.
- [58] P. Wojtowicz, H. Janečková, D. Friedecký, T. Adam, Techniky metabolomiky v biomedicíně, *Chem. List*. 107 (2013) 3–11.
- [59] S. Hayton, G.L. Maker, I. Mullaney, R.D. Trengove, Experimental design and reporting standards for metabolomics studies of mammalian cell lines, *Cell. Mol. Life Sci.* 74 (2017) 4421–4441. doi:10.1007/s00018-017-2582-1.
- [60] A. Zhang, H. Sun, X. Wu, X. Wang, Urine metabolomics, *Clin. Chim. Acta*. 414 (2012) 65–69. doi:10.1016/j.cca.2012.08.016.
- [61] L. Yi, S. Shi, Z. Yi, R. He, H. Lu, Y. Liang, MeOx-TMS derivatization for GC-MS metabolic profiling of urine and application in the discrimination between normal C57BL/6J and type 2 diabetic KK-Ay mice, *Anal. Methods*. 6 (2014) 4380–4387. doi:10.1039/C3AY41522H.
- [62] O. Fiehn, J. Kopka, R.N. Trethewey, L. Willmitzer, Identification of Uncommon Plant Metabolites Based on Calculation of Elemental Compositions Using Gas Chromatography and Quadrupole Mass Spectrometry, *Anal. Chem.* 72 (2000) 3573–3580. doi:10.1021/ac991142i.
- [63] B.-J. Webb-Robertson, Y.-M. Kim, E.M. Zink, K.A. Hallaian, Q. Zhang, R. Madupu, K.M. Waters, T.O. Metz, A Statistical Analysis of the Effects of Urease Pre-treatment on the Measurement of the Urinary Metabolome by Gas Chromatography-Mass Spectrometry., *Metabolomics*. 10 (2014) 897–908. doi:10.1007/s11306-014-0642-1.
- [64] J.D. Shoemaker, W.H. Elliott, Automated screening of urine samples for carbohydrates, organic and amino acids after treatment with urease, *J. Chromatogr. B Biomed. Sci. Appl.* 562 (1991) 125–138. doi:10.1016/0378-4347(91)80571-S.

- [65] M.K. Das, R. Arya, S. Debnath, R. Debnath, A. Lodh, S.C. Bishwal, A. Das, R.K. Nanda, Global Urine Metabolomics in Patients Treated with First-Line Tuberculosis Drugs and Identification of a Novel Metabolite of Ethambutol, *Antimicrob. Agents Chemother.* 60 (2016) 2257–2264. doi:10.1128/AAC.02586-15.
- [66] E.C.Y. Chan, K.K. Pasikanti, J.K. Nicholson, Global urinary metabolic profiling procedures using gas chromatography–mass spectrometry, *Nat. Protoc.* 6 (2011) 1483–1499. doi:10.1038/nprot.2011.375.
- [67] P. Zheng, Y. Wang, H. Lu, X. Zhou, T. Tang, R. Fan, C. Zhang, H. Cui, Y. Wang, J. Luo, Plasma Metabolomics Analysis Based on GC-MS in Infertile Males with Kidney-Yang Deficiency Syndrome, 2017 (2017).
- [68] D. Feng, Z. Xia, J. Zhou, H. Lu, C. Zhang, R. Fan, S. Di, G. Cao, Metabolomics reveals the effect of Xuefu Zhuyu Decoction on plasma metabolism in rats with acute traumatic brain injury, 8 (2017) 94692–94710.
- [69] B. Cao, J. Aa, G. Wang, X. Wu, L. Liu, M. Li, J. Shi, X. Wang, C. Zhao, T. Zheng, S. Guo, J. Duan, GC–TOFMS analysis of metabolites in adherent MDCK cells and a novel strategy for identifying intracellular metabolic markers for use as cell amount indicators in data normalization, *Anal. Bioanal. Chem.* 400 (2011) 2983–2993. doi:10.1007/s00216-011-4981-8.
- [70] O. Fiehn, Metabolomics by Gas Chromatography-Mass Spectrometry: Combined Targeted and Untargeted Profiling, in: *Curr. Protoc. Mol. Biol.*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2016: pp. 30.4.1-30.4.32. doi:10.1002/0471142727.mb3004s114.
- [71] B. Berger, J. Peng, M. Singh, Computational solutions for omics data, *Nat. Rev. Genet.* 14 (2013) 333–346. doi:10.1038/nrg3433.
- [72] R. Di Guida, J. Engel, J.W. Allwood, R.J.M. Weber, M.R. Jones, U. Sommer, M.R. Viant, W.B. Dunn, Non-targeted UHPLC-MS metabolomic data processing methods: a comparative investigation of normalisation, missing value imputation, transformation and scaling, *Metabolomics.* 12 (2016) 1–14. doi:10.1007/s11306-016-1030-9.
- [73] O. Fiehn, Metabolomics by Gas Chromatography-Mass Spectrometry: Combined Targeted and Untargeted Profiling, in: *Curr. Protoc. Mol. Biol.*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2016: pp. 30.4.1-30.4.32. doi:10.1002/0471142727.mb3004s114.
- [74] W.B. Dunn, I.D. Wilson, A.W. Nicholls, D. Broadhurst, The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans, *Bioanalysis.* 4 (2012) 2249–2264. doi:10.4155/bio.12.204.

- [75] J. Aitchison, *The Statistical Analysis of Compositional Data*, Chapman & Hall, Ltd., London, UK, UK, 1986.
- [76] A. Kalivodová, K. Hron, P. Filzmoser, L. Najdekr, H. Janečková, T. Adam, PLS-DA for compositional data with application to metabolomics, *J. Chemom.* 29 (2015) 21–28. doi:10.1002/cem.2657.
- [77] L. Najdekr, A. Gardlo, L. Mádrová, D. Friedecký, H. Janečková, E.S. Correa, R. Goodacre, T. Adam, Oxidized phosphatidylcholines suggest oxidative stress in patients with medium-chain acyl-CoA dehydrogenase deficiency, *Talanta*. 139 (2015) 62–66. doi:10.1016/j.talanta.2015.02.041.
- [78] F. Dieterle, A. Ross, G. Schlotterbeck, H. Senn, Probabilistic Quotient Normalization as Robust Method to Account for Dilution of Complex Biological Mixtures. Application in <sup>1</sup>H NMR Metabonomics, *Anal. Chem.* 78 (2006) 4281–4290. doi:10.1021/ac051632c.
- [79] R.A. van den Berg, H.C.J. Hoefsloot, J.A. Westerhuis, A.K. Smilde, M.J. van der Werf, Centering, scaling, and transformations: improving the biological information content of metabolomics data., *BMC Genomics*. 7 (2006) 142. doi:10.1186/1471-2164-7-142.
- [80] T. Vu, E. Riekeberg, Y. Qiu, R. Powers, Comparing normalization methods and the impact of noise, *Metabolomics*. 14 (2018) 108. doi:10.1007/s11306-018-1400-6.
- [81] M. MELOUN, J. MILITKÝ, *Statistické zpracování experimentálních dat v chemometrii, biometrii, ekonometrii a v dalších oborech přírodních, technických a společenských věd*, Praha: East Publishing, 1998.
- [82] J. Lever, M. Krzywinski, N. Altman, Principal component analysis, *Nat. Methods*. 14 (2017) 641–642. doi:10.1038/nmeth.4346.
- [83] P. Jonsson, J. Gullberg, A. Nordström, M. Kusano, M. Kowalczyk, M. Sjöström, T. Moritz, A strategy for identifying differences in large series of metabolomic samples analyzed by GC/MS., *Anal. Chem.* 76 (2004) 1738–45. doi:10.1021/ac0352427.
- [84] M. Barker, W. Rayens, Partial least squares for discrimination, *J. Chemom.* 17 (2003) 166–173. doi:10.1002/cem.785.
- [85] T.N. Tran, N.L. Afanador, L.M.C. Buydens, L. Blanchet, Interpretation of variable importance in Partial Least Squares with Significance Multivariate Correlation (sMC), *Chemom. Intell. Lab. Syst.* 138 (2014) 153–160. doi:10.1016/j.chemolab.2014.08.005.
- [86] H.-W. Cho, S.B. Kim, M.K. Jeong, Y. Park, N.G. Miller, T.R. Ziegler, D.P. Jones, Discovery of metabolite features for the modelling and analysis of high-resolution NMR spectra., *Int. J. Data Min. Bioinform.* 2 (2008) 176–92. <http://www.ncbi.nlm.nih.gov/pubmed/18767354>  
<http://www.pubmedcentral.nih.gov/arti>

clerender.fcgi?artid=PMC3883573.

- [87] P. Banerjee, S. Ghosh, M. Dutta, E. Subramani, J. Khalpada, S. RoyChoudhury, B. Chakravarty, K. Chaudhury, Identification of Key Contributory Factors Responsible for Vascular Dysfunction in Idiopathic Recurrent Spontaneous Miscarriage, *PLoS One*. 8 (2013) e80940. doi:10.1371/journal.pone.0080940.
- [88] N. Akarachantachote, S. Chadcham, K. Saithanu, CUTOFF THRESHOLD OF VARIABLE IMPORTANCE IN PROJECTION FOR VARIABLE SELECTION, *Int. J. Pure Applied Math*. 94 (2014) 307–322. doi:10.12732/ijpam.v94i3.2.
- [89] M. Pérez-Enciso, M. Tenenhaus, Prediction of clinical outcome with microarray data: a partial least squares discriminant analysis (PLS-DA) approach., *Hum. Genet*. 112 (2003) 581–92. doi:10.1007/s00439-003-0921-9.
- [90] J. Podani, I. Miklos, Resemblance Coefficients and the Horseshoe Effect in Principal Coordinates Analysis, *Ecology*. 83 (2002) 3331. doi:10.2307/3072083.
- [91] J.T. Morton, L. Toran, A. Edlund, J.L. Metcalf, C. Lauber, R. Knight, Uncovering the Horseshoe Effect in Microbial Analyses, *MSystems*. 2 (2017) 1–7. doi:10.1128/mSystems.00166-16.
- [92] W.W. Wells, T. Chin, B. Weber, Quantitative analysis of serum and urine sugars by gas chromatography, *Clin. Chim. Acta*. 10 (1964) 352–359. doi:10.1016/0009-8981(64)90066-X.
- [93] I. Matsumoto, T. Kuhara, A new chemical diagnostic method for inborn errors of metabolism by mass spectrometry—rapid, practical, and simultaneous urinary metabolites analysis, *Mass Spectrom. Rev*. 15 (1996) 43–57. doi:10.1002/(SICI)1098-2787(1996)15:1<43::AID-MAS3>3.0.CO;2-B.
- [94] T. Kuhara, Diagnosis of inborn errors of metabolism using filter paper urine, urease treatment, isotope dilution and gas chromatography-mass spectrometry, *J Chromatogr B Biomed Sci Appl*. 758 (2001) 3–25.
- [95] D.D. van Slyke, G.E. Cullen, Mode of Action of Urease, *J. Biol. Chem*. 19 (1914) 141–180.
- [96] J.F. Marlier, L.I. Robins, K.A. Tucker, J. Rawlings, M.A. Anderson, W.W. Cleland, A Kinetic and Isotope Effect Investigation of the Urease-Catalyzed Hydrolysis of Hydroxyurea, *Biochemistry*. 49 (2010) 8213–8219. doi:10.1021/bi100890v.
- [97] R. Karlíková, J. Šíroková, D. Friedecký, E. Faber, M. Hrdá, K. Mičová, I. Fikarová, A. Gardlo, H. Janečková, I. Vrobel, T. Adam, Metabolite Profiling of the Plasma and Leukocytes of Chronic Myeloid Leukemia Patients, *J. Proteome Res*. 15 (2016) 3158–3166. doi:10.1021/acs.jproteome.6b00356.

- [98] T. Kind, V. Tolstikov, O. Fiehn, R.H. Weiss, A comprehensive urinary metabolomic approach for identifying kidney cancer, *Anal. Biochem.* 363 (2007) 185–195. doi:10.1016/j.ab.2007.01.028.
- [99] M. Jořenek, *Studium biologicky aktivních látek imobilizovaných na magnetických nosičích a jejich použití v biotechnologických procesech*, Univerzita Palackého v Olomouci, 2016.
- [100] K. Grob, S. Rennhard, Evaluation of syringe handling techniques for injections into vaporizing GC injectors, *J. High Resolut. Chromatogr.* 3 (1980) 627–633. doi:10.1002/jhrc.1240031205.
- [101] R. Tautenhahn, G.J. Patti, D. Rinehart, G. Siuzdak, XCMS Online: a web-based platform to process untargeted metabolomic data., *Anal. Chem.* 84 (2012) 5035–5039. doi:10.1021/ac300698c.
- [102] C.A. Smith, E.J. Want, G. O’Maille, R. Abagyan, G. Siuzdak, XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification, *Anal. Chem.* 78 (2006) 779–787. doi:10.1021/ac051437y.
- [103] A. Gardlo, AlzbetaG/Metabol: First version, (2019). doi:10.5281/zenodo.3235775.
- [104] S.U. Bajad, W. Lu, E.H. Kimball, J. Yuan, C. Peterson, J.D. Rabinowitz, Separation and quantitation of water soluble cellular metabolites by hydrophilic interaction chromatography-tandem mass spectrometry, *J. Chromatogr. A.* 1125 (2006) 76–88. doi:10.1016/j.chroma.2006.05.019.
- [105] J.L. Little, Artifacts in trimethylsilyl derivatization reactions and ways to avoid them, *J. Chromatogr. A.* 844 (1999) 1–22. doi:10.1016/S0021-9673(99)00267-8.
- [106] R.S. Clements, W.R. Starnes, An improved method for the determination of urinary myoinositol by gas-liquid chromatography, *Biochem. Med.* 12 (1975) 200–204. doi:10.1016/0006-2944(75)90112-X.
- [107] F. Palmas, M. Mussap, C. Fattuoni, Urine metabolome analysis by gas chromatography–mass spectrometry (GC–MS): Standardization and optimization of protocols for urea removal and short-term sample storage, *Clin. Chim. Acta.* 485 (2018) 236–242. doi:10.1016/j.cca.2018.07.006.
- [108] D.C. Lehotay, J.T.R. Clarke, P. Renaldo, Organic Acidurias and Related Abnormalities, *Crit. Rev. Clin. Lab. Sci.* 32 (1995) 377–429. doi:10.3109/10408369509084689.
- [109] O. Fiehn, G. Wohlgemuth, M. Scholz, T. Kind, D.Y. Lee, Y. Lu, S. Moon, B. Nikolau, Quality control for plant metabolomics: reporting MSI-compliant studies, *Plant J.* 53 (2008) 691–704. doi:10.1111/j.1365-313X.2007.03387.x.
- [110] A. Liu, M.M. Kushnir, W.L. Roberts, M. Pasquali, Solid phase extraction procedure for urinary organic acid analysis by gas chromatography mass spectrometry., *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* 806 (2004) 283–7. doi:10.1016/j.jchromb.2004.03.048.

- [111] Y. Koh, K.K. Pasikanti, C.W. Yap, E.C.Y. Chan, Comparative evaluation of software for retention time alignment of gas chromatography/time-of-flight mass spectrometry-based metabonomic data, *J. Chromatogr. A.* 1217 (2010) 8308–8316. doi:10.1016/j.chroma.2010.10.101.
- [112] G.F. Hoffmann, F.K. Trefz, P.G. Barth, H.J. Böhles, B. Biggemann, H.J. Bremer, E. Christensen, M. Frosch, F. Hanefeld, D.H. Hunneman, Glutaryl-coenzyme A dehydrogenase deficiency: a distinct encephalopathy., *Pediatrics.* 88 (1991) 1194–203. <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed2&NEWS=N&AN=1992029708>.
- [113] G.L. Hedlund, N. Longo, M. Pasquali, Glutaric acidemia type 1., *Am. J. Med. Genet. C. Semin. Med. Genet.* 142C (2006) 86–94. doi:10.1002/ajmg.c.30088.
- [114] P.J. Slonecker, X. Li, T.H. Ridgway, J.G. Dorsey, Informational orthogonality of two-dimensional chromatographic separations., *Anal. Chem.* 68 (1996) 682–9. doi:10.1021/ac950852v.
- [115] J.M. Davis, D.R. Stoll, P.W. Carr, Effect of first-dimension undersampling on effective peak capacity in comprehensive two-dimensional separations., *Anal. Chem.* 80 (2008) 461–73. doi:10.1021/ac071504j.
- [116] P. Wojtowicz, J. Zrostlíková, T. Kovalczuk, J. Schůrek, T. Adam, Evaluation of comprehensive two-dimensional gas chromatography coupled to time-of-flight mass spectrometry for the diagnosis of inherited metabolic disorders using an automated data processing strategy., *J. Chromatogr. A.* 1217 (2010) 8054–61. doi:10.1016/j.chroma.2010.09.067.
- [117] E.A. McGaw, K.W. Phinney, M.S. Lowenthal, Comparison of orthogonal liquid and gas chromatography–mass spectrometry platforms for the determination of amino acid concentrations in human plasma, *J. Chromatogr. A.* 1217 (2010) 5822–5831. doi:10.1016/j.chroma.2010.07.025.
- [118] D. Ryan, P. Morrison, P. Marriott, Orthogonality considerations in comprehensive two-dimensional gas chromatography., *J. Chromatogr. A.* 1071 (2005) 47–53. doi:10.1016/j.chroma.2004.09.020.
- [119] M. Gilar, J. Fridrich, M.R. Schure, A. Jaworski, Comparison of orthogonality estimation methods for the two-dimensional separations of peptides., *Anal. Chem.* 84 (2012) 8722–32. doi:10.1021/ac3020214.
- [120] D. Li, C. Jakob, O. Schmitz, Practical considerations in comprehensive two-dimensional liquid chromatography systems (LCxLC) with reversed-phases in both dimensions., *Anal. Bioanal. Chem.* 407 (2015) 153–67. doi:10.1007/s00216-014-8179-8.



- [121] M.R. Schure, J.M. Davis, Orthogonal separations: Comparison of orthogonality metrics by statistical analysis, *J. Chromatogr. A.* 1414 (2015) 60–76. doi:10.1016/j.chroma.2015.08.029.
- [122] D. Haas, G.F. Hoffmann, Mevalonate kinase deficiencies: from mevalonic aciduria to hyperimmunoglobulinemia D syndrome., *Orphanet J. Rare Dis.* 1 (2006) 13. doi:10.1186/1750-1172-1-13.
- [123] B. Omais, M. Courtiade, N. Charon, J. Ponthus, D. Thiébaud, Considerations on orthogonality duality in comprehensive two-dimensional gas chromatography., *Anal. Chem.* 83 (2011) 7550–4. doi:10.1021/ac201103e.
- [124] D.N. Bassanese, B.J. Holland, X.A. Conlan, P.S. Francis, N.W. Barnett, P.G. Stevenson, Protocols for finding the most orthogonal dimensions for two-dimensional high performance liquid chromatography., *Talanta.* 134 (2015) 402–8. doi:10.1016/j.talanta.2014.11.037.
- [125] W. Hoeffding, Scale—Invariant Correlation Theory, in: 1994: pp. 57–107. doi:10.1007/978-1-4612-0865-5\_4.
- [126] M.R. Schure, The dimensionality of chromatographic separations, *J. Chromatogr. A.* 1218 (2011) 293–302. doi:10.1016/j.chroma.2010.11.016.
- [127] J. V Seeley, S.K. Seeley, E.K. Libby, J.D. McCurry, Analysis of biodiesel/petroleum diesel blends with comprehensive two-dimensional gas chromatography, *J. Chromatogr. Sci.* 45 (2007) 650–656.
- [128] C. Manzano, E. Hoh, S.L.M. Simonich, Improved Separation of Complex Polycyclic Aromatic Hydrocarbon Mixtures Using Novel Column Combinations in GC × GC/ToF-MS, *Environ. Sci. Technol.* 46 (2012) 7677–7684. doi:10.1021/es301790h.
- [129] J.-F. Focant, A. Sjödin, D.G. Patterson, Improved separation of the 209 polychlorinated biphenyl congeners using comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry, *J. Chromatogr. A.* 1040 (2004) 227–238. doi:10.1016/j.chroma.2004.04.003.
- [130] J. Leonhardt, T. Teutenberg, G. Buschmann, O. Gassner, T.C. Schmidt, A new method for the determination of peak distribution across a two-dimensional separation space for the identification of optimal column combinations, *Anal. Bioanal. Chem.* 408 (2016). doi:10.1007/s00216-016-9911-3.
- [131] N.P. Vasquez, M. Crosnier de bellaistre-Bonose, N. Lévêque, E. Thioulouse, D. Doummar, T. Billette de Villemeur, D. Rodriguez, R. Couderc, S. Robin, C. Courderot-Masuyer, F. Moussa, Advances in the metabolic profiling of acidic compounds in children’s urines achieved by comprehensive two-dimensional gas chromatography, *J. Chromatogr. B.* 1002 (2015) 130–138.

doi:10.1016/j.jchromb.2015.08.006.

- [132] J. Vial, B. Pezous, D. Thiébaud, P. Sassiati, B. Teillet, X. Cahours, I. Rivals, The discriminant pixel approach: A new tool for the rational interpretation of GCxGC-MS chromatograms, *Talanta*. 83 (2011) 1295–1301. doi:10.1016/j.talanta.2010.07.059.
- [133] C. Cordero, P. Rubiolo, B. Sgorbini, M. Galli, C. Bicchi, Comprehensive two-dimensional gas chromatography in the analysis of volatile samples of natural origin: a multidisciplinary approach to evaluate the influence of second dimension column coated with mixed stationary phases on system orthogonality., *J. Chromatogr. A*. 1132 (2006) 268–79. doi:10.1016/j.chroma.2006.07.067.
- [134] L.M. Blumberg, *Comprehensive Chromatography in Combination with Mass Spectrometry*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2011. doi:10.1002/9781118003466.
- [135] J. Krupčík, J. Garaj, P. Čellár, G. Guiochon, Calculation of the peak capacity in capillary gas chromatography, *J. Chromatogr. A*. 312 (1984) 1–10. doi:10.1016/S0021-9673(01)92759-1.
- [136] L.M. Blumberg, Linear peak capacity of a comprehensive multi-dimensional separation, *J. Sep. Sci.* 31 (2008) 3352–3357. doi:10.1002/jssc.200800244.
- [137] M.S. Klee, J. Cochran, M. Merrick, L.M. Blumberg, Evaluation of conditions of comprehensive two-dimensional gas chromatography that yield a near-theoretical maximum in peak capacity gain, *J. Chromatogr. A*. 1383 (2015) 151–159. doi:10.1016/j.chroma.2015.01.031.
- [138] R.C.Y. Ong, P.J. Marriott, A Review of Basic Concepts in Comprehensive Two-Dimensional Gas Chromatography, *J. Chromatogr. Sci.* 40 (2002) 276–291. doi:10.1093/chromsci/40.5.276.
- [139] K. Kalachova, J. Pulkrabova, T. Cajka, L. Drabova, J. Hajslova, Implementation of comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry for the simultaneous determination of halogenated contaminants and polycyclic aromatic hydrocarbons in fish, *Anal. Bioanal. Chem.* 403 (2012) 2813–2824. doi:10.1007/s00216-012-6095-3.
- [140] M.M. Koek, B. Muilwijk, L.L.P. van Stee, T. Hankemeier, Higher mass loadability in comprehensive two-dimensional gas chromatography–mass spectrometry for improved analytical performance in metabolomics analysis, *J. Chromatogr. A*. 1186 (2008) 420–429. doi:10.1016/j.chroma.2007.11.107.
- [141] J. Beens, J. Blomberg, P.J. Schoenmakers, Proper Tuning of Comprehensive Two-Dimensional Gas Chromatography (GCxGC) to Optimize the Separation of Complex Oil Fractions, *J. High Resolut. Chromatogr.* 23 (2000) 182–188. doi:10.1002/(SICI)1521-4168(20000301)23:3<182::AID-JHRC182>3.3.CO;2-5.

- [142] W. Engewald, K. Dettmer-Wilde, H. Rotzsche, Columns and Stationary Phases, in: Pract. Gas Chromatogr., Springer Berlin Heidelberg, Berlin, Heidelberg, 2014: pp. 59–116. doi:10.1007/978-3-642-54640-2\_3.
- [143] B.S. Mitrevski, J.T. Brenna, Y. Zhang, P.J. Marriott, Application of comprehensive two-dimensional gas chromatography to sterols analysis, *J. Chromatogr. A.* 1214 (2008) 134–142. doi:10.1016/j.chroma.2008.10.045.
- [144] M.M. Koek, F.M. van der Kloet, R. Kleemann, T. Kooistra, E.R. Verheij, T. Hankemeier, Semi-automated non-target processing in GC × GC–MS metabolomics analysis: applicability for biomedical studies, *Metabolomics.* 7 (2011) 1–14. doi:10.1007/s11306-010-0219-6.
- [145] J.C. Schoeman, I. du Preez, D.T. Loots, A comparison of four sputum pre-extraction preparation methods for identifying and characterising *Mycobacterium tuberculosis* using GCxGC-TOFMS metabolomics, *J. Microbiol. Methods.* 91 (2012) 301–311. doi:10.1016/j.mimet.2012.09.002.
- [146] N. Koen, S.V. van Breda, D.T. Loots, Elucidating the antimicrobial mechanisms of colistin sulfate on *Mycobacterium tuberculosis* using metabolomics, *Tuberculosis.* 111 (2018) 14–19. doi:10.1016/j.tube.2018.05.001.
- [147] S. KC, M. Liu, Q. Zhang, K. Fan, Y. Shi, J. Ruan, Metabolic Changes of Amino Acids and Flavonoids in Tea Plants in Response to Inorganic Phosphate Limitation, *Int. J. Mol. Sci.* 19 (2018) 3683. doi:10.3390/ijms19113683.
- [148] A.M. Miguez, M.P. McNerney, M.P. Styczynski, Metabolomics Analysis of the Toxic Effects of the Production of Lycopene and Its Precursors, *Front. Microbiol.* 9 (2018) 1–15. doi:10.3389/fmicb.2018.00760.
- [149] D. Bressanello, E. Liberto, M. Collino, F. Chiazza, R. Mastrocola, S.E. Reichenbach, C. Bicchi, C. Cordero, Combined untargeted and targeted fingerprinting by comprehensive two-dimensional gas chromatography: revealing fructose-induced changes in mice urinary metabolic signatures, *Anal. Bioanal. Chem.* 410 (2018) 2723–2737. doi:10.1007/s00216-018-0950-9.
- [150] T. Miyazaki, K. Okada, T. Yamashita, M. Miyazaki, Two-dimensional gas chromatography time-of-flight mass spectrometry-based serum metabolic fingerprints of neonatal calves before and after first colostrum ingestion, *J. Dairy Sci.* 100 (2017) 4354–4364. doi:10.3168/jds.2017-12557.
- [151] Y. Izadmanesh, E. Garreta-Lara, J.B. Ghasemi, S. Lacorte, V. Matamoros, R. Tauler, Chemometric analysis of comprehensive two dimensional gas chromatography–mass spectrometry metabolomics data, *J. Chromatogr. A.* 1488 (2017) 113–125. doi:10.1016/j.chroma.2017.01.052.

- [152] Y. Liu, X. Wang, Y. Li, X. Chen, Metabolomic analysis of short-term sulfamethazine exposure on marine medaka ( *Oryzias melastigma* ) by comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry, *Aquat. Toxicol.* 198 (2018) 269–275. doi:10.1016/j.aquatox.2018.03.006.
- [153] D.T. Loots, C.C. Swanepoel, M. Newton-Foot, N.C. Gey van Pittius, A metabolomics investigation of the function of the ESX-1 gene cluster in mycobacteria, *Microb. Pathog.* 100 (2016) 268–275. doi:10.1016/j.micpath.2016.10.008.
- [154] C.H. Weinert, B. Egert, S.E. Kulling, On the applicability of comprehensive two-dimensional gas chromatography combined with a fast-scanning quadrupole mass spectrometer for untargeted large-scale metabolomics, *J. Chromatogr. A.* 1405 (2015) 156–167. doi:10.1016/j.chroma.2015.04.011.
- [155] R.E. Mohler, K.M. Dombek, J.C. Hoggard, E.T. Young, R.E. Synovec, Comprehensive two-dimensional gas chromatography time-of-flight mass spectrometry analysis of metabolites in fermenting and respiring yeast cells., *Anal. Chem.* 78 (2006) 2700–9. doi:10.1021/ac052106o.
- [156] L. Sweetman, P. Ashcraft, J. Bennett-Firmin, Quantitative Organic Acids in Urine by Two Dimensional Gas Chromatography-Time of Flight Mass Spectrometry (GCxGC-TOFMS), in: U. Garg (Ed.), Springer New York, New York, NY, 2016: pp. 183–197. doi:10.1007/978-1-4939-3182-8\_20.
- [157] B. Egert, C.H. Weinert, S.E. Kulling, A peaklet-based generic strategy for the untargeted analysis of comprehensive two-dimensional gas chromatography mass spectrometry data sets, *J. Chromatogr. A.* 1405 (2015) 168–177. doi:10.1016/j.chroma.2015.05.056.
- [158] J.H. Winnike, X. Wei, K.J. Knagge, S.D. Colman, S.G. Gregory, X. Zhang, Comparison of GC-MS and GCxGC-MS in the analysis of human serum samples for biomarker discovery., *J. Proteome Res.* 14 (2015) 1810–7. doi:10.1021/pr5011923.
- [159] S. Ly-Verdú, T.M. Gröger, J.M. Arteaga-Salas, S. Brandmaier, M. Kahle, S. Neschen, M. Harbě de Angelis, R. Zimmermann, Combining metabolomic non-targeted GCxGC-ToF-MS analysis and chemometric ASCA-based study of variances to assess dietary influence on type 2 diabetes development in a mouse model, *Anal. Bioanal. Chem.* 407 (2015) 343–354. doi:10.1007/s00216-014-8227-4.
- [160] K.K. Pasikanti, K. Esuvaranathan, Y. Hong, P.C. Ho, R. Mahendran, L. Raman Nee Mani, E. Chiong, E.C.Y. Chan, Urinary Metabotyping of Bladder Cancer Using Two-Dimensional Gas Chromatography Time-of-Flight Mass Spectrometry, *J. Proteome Res.* 12 (2013) 3865–3873.

doi:10.1021/pr4000448.

- [161] K.W. Phinney, G. Ballihaut, M. Bedner, B.S. Benford, J.E. Camara, S.J. Christopher, W.C. Davis, N.G. Dodder, G. Eppe, B.E. Lang, S.E. Long, M.S. Lowenthal, E.A. McGaw, K.E. Murphy, B.C. Nelson, J.L. Prendergast, J.L. Reiner, C.A. Rimmer, L.C. Sander, M.M. Schantz, K.E. Sharpless, L.T. Sniegowski, S.S.-C. Tai, J.B. Thomas, T.W. Vetter, M.J. Welch, S.A. Wise, L.J. Wood, W.F. Guthrie, C.R. Hagwood, S.D. Leigh, J.H. Yen, N.-F. Zhang, M. Chaudhary-Webb, H. Chen, Z. Fazili, D.J. LaVoie, L.F. McCoy, S.S. Momin, N. Paladugula, E.C. Pendergrast, C.M. Pfeiffer, C.D. Powers, D. Rabinowitz, M.E. Rybak, R.L. Schleicher, B.M.H. Toombs, M. Xu, M. Zhang, A.L. Castle, Development of a Standard Reference Material for metabolomics research., *Anal. Chem.* 85 (2013) 11732–8. doi:10.1021/ac402689t.
- [162] M. Adahchour, J. Beens, R.J.J. Vreuls, U.A.T. Brinkman, Recent developments in comprehensive two-dimensional gas chromatography (GC×GC), *TrAC Trends Anal. Chem.* 25 (2006) 438–454. doi:10.1016/j.trac.2006.03.002.
- [163] S. Whitmarsh, I. Analysis, Ionic Liquid Stationary Phases : Application in Gas Chromatographic Analysis of, *Chromatogr. Today.* (2012) 12–15. <https://www.chromatographytoday.com/article/gc-mdgc/32/unassigned-independent-article/ionic-liquid-stationary-phases-application-in-gas-chromatographic-analysis-of-nbsp-polar-components-of-fuels-and-lubricants/1132>.
- [164] N. Lenca, C.F. Poole, A system map for the ionic liquid stationary phase 1,12-di(triethylphosphonium)dodecane bis(trifluoromethylsulfonyl)imide for gas chromatography, *J. Chromatogr. A.* 1525 (2017) 138–144. doi:10.1016/j.chroma.2017.10.015.
- [165] A.X. Zeng, S.-T. Chin, Y. Nolvachai, C. Kulsing, L.M. Sidisky, P.J. Marriott, Characterisation of capillary ionic liquid columns for gas chromatography–mass spectrometry analysis of fatty acid methyl esters, *Anal. Chim. Acta.* 803 (2013) 166–173. doi:10.1016/j.aca.2013.07.002.
- [166] K. Pospiskova, I. Safarik, Low-cost, easy-to-prepare magnetic chitosan microparticles for enzymes immobilization, *Carbohydr. Polym.* 96 (2013) 545–548. doi:10.1016/j.carbpol.2013.04.014.
- [167] W. Nowik, M. Bonose, S. Héron, M. Nowik, A. Tchaplá, Assessment of Two-Dimensional Separative Systems Using the Nearest Neighbor Distances Approach. Part 2: Separation Quality Aspects, *Anal. Chem.* 85 (2013) 9459–9468. doi:10.1021/ac4012717.

## 15 Seznam zkratk

1D	jednodimenzionální
$^1t_r$	retenční čas první dimenze
2D	dvojdimenzionální
$^2t_r$	retenční čas druhé dimenze
AE	metoda hvězdičkových rovnic
AN	aritmetický průměr vzdáleností nejbližších sousedů
BSTFA	bis(trimethylsilyl)trifluoroacetamid + 1% chlortrimethylsilan
COA	kontrolní materiál na bázi lidské moči s přidavkem některých metabolitů
CV	průměrná relativní směrodatná odchylka
D1	první dimenze
D2	druhá dimenze
dTIC	část TIC, která po dekonvoluci náleží jednomu z koelujících analytů
ECD	detektor elektronového záchytu
EtOx	O-ethylhydroxylamin hydrochlorid
EX	vzorky připravené pomocí kyselé extrakce do ethylacetátu
FID	plamenový ionizační detektor
FUT	vzorky připravené pomocí volné ureasy
GC	plynová chromatografie
GC×GC	dvojdimenzionální plynová chromatografie
HN	harmonický průměr vzdáleností nejbližších sousedů
CH	konvexní obálka
IS	interní standard
IUT	vzorky připravené pomocí imobilizované ureasy
LC	kapalinová chromatografie
MA	přístup založený na modelování
MS	hmotnostní spektrometrie
MSTFA	N-trimethylsilyl-N-methyl trifluoroacetamid

NMR	nukleární magnetická rezonance
NN	metoda nejbližšího souseda
NT	vzorky bez úpravy
O	ortogonalita
P	počet píků
PC	Pearsonova korelace
PC1	první hlavní komponenta
PC2	druhá hlavní komponenta
PCA	analýza hlavních komponent
PLS-DA	diskriminační analýza na základě metody nejmenších čtverců
QC	vzorek pro kontrolu kvality
r	korelační koeficient
r <sup>2</sup>	koeficient determinace
RI	retenční index
SC	geometrické pokrytí separačního povrchu
SRM	monitorování vybrané reakce
t <sub>0</sub>	mrtvý čas
TIC	celkový iontový proud
TMCS	chlortrimethylsilan
TOF MS	hmotnostní spektrometr obsahující hmotnostní analyzátor doby letu
t <sub>r</sub>	retenční čas
VIP	graf významnosti proměnných v projekci použité pro PLS-DA model

## 16 Doplňující informace

Doplňující informace obsahují tabulky (Tab. S1-S4), které jsou příliš dlouhé na to, aby byly vloženy do textu práce:

Tab. S1: Označení, retenční čas (v IUT) a poměry mediánů analytů pro jednotlivé přístupy přípravy vzorku analyzovaných pomocí LC-MS

Tab. S2: Průměrné hodnoty ortogonalit jednotlivých typů modelů se 100 body získané různými přístupy výpočtu ortogonalit  $\pm 2SD$  pro 50 modelů stejného typu. Hodnoty jsou uvedeny v %, s výjimkou AN a HN, jejichž výsledkem jsou průměrné nejmenší vzdálenosti mezi nejbližšími sousedy.

Tab. S3: Přehled syntetických chromatogramů nízkodimenzionálních modelů typu E a hodnot ortogonalit [%] získaných prostřednictvím jednotlivých deskriptorů.

Tab. S4: Přehled syntetických chromatogramů vysokodimenzionálních modelů typu E a hodnot ortogonalit [%] získaných prostřednictvím jednotlivých deskriptorů.



Tab. S1: Označení, retenční čas (v IUT) a poměry mediánů analytů pro jednotlivé přístupy přípravy vzorku analyzovaných pomocí LC-MS

Označení	t <sub>r</sub> [min]	Analyt	FUT/IUT	NT/IUT
C001	6,28	Glutamát	0,99	1,00
C002	7,23	2-oxoglutarát	1,13	1,15
C003	3,69	Hexosy – glukosa, fruktosa, galaktosa, manosa	0,87	0,98
C004	7,14	Oxaloacetát/glutarát/ethylmalonát	1,07	1,05
C005	4,57	Glycin	1,30	1,26
C006	4,35	Alanin/sarkosin/beta-alanin	1,01	1,04
C007	7,13	Sukcinát/methylmalonát	1,03	0,98
C008	5,16	Lysin	1,03	1,19
C009	7,10	Aspartát	1,06	1,03
C010	4,53	Glutamin	0,98	1,12
C011	4,74	Serin	1,09	1,14
C012	4,13	Methionin	1,04	1,12
C013	4,14	Tryptofan	0,89	1,21
C014	4,06	Fenylalanin	0,99	1,08
C015	4,48	Tyrosin	1,15	1,17
C016	7,66	Glycerol-3-fosfát	1,04	0,99
C017	2,37	Uracil	1,05	3,03
C018	4,31	Cholin	0,93	0,94
C019	4,08	Leucin	1,20	1,22
C020	3,85	Isoleucin/alloisoleucin/norleucin/hydroxyprolin	1,05	1,06
C021	7,22	IMP	0,99	1,08
C022	4,80	Histidine	1,20	1,22
C023	4,13	Laktosa	37,07	0,64
C024	3,46	N-acetyl- glukosamin/galaktosamin/manosamine	1,10	1,11
C025	3,41	Adenin	1,09	0,80
C026	4,72	Asparagin	1,20	1,08
C027	4,48	Homocystein	1,08	1,07
C028	8,27	Citrát	1,11	1,00

Tab. S1: Pokračování

Označení	t <sub>r</sub> [min]	Analyt	FUT/IUT	NT/IUT
C029	8,27	Isocitrát	1,04	1,03
C030	4,16	Valin	1,09	1,18
C031	4,54	Threonin/homoserin	1,07	1,12
C032	6,20	Glukuronát	1,06	0,62
C033	3,17	Adenosin	1,02	1,12
C034	4,45	Taurin	1,24	1,27
C035	3,12	Riboflavin	0,82	1,30
C036	5,86	Mevalonolakton	0,98	1,10
C037	4,19	Hypoxanthin	1,01	1,18
C038	5,32	Orotát	1,05	0,95
C039	3,16	Uridin	1,14	1,09
C040	4,08	5-aminolevulinát/Kreatin	1,07	1,14
C041	4,06	Kynurenin	1,30	1,38
C042	3,84	Glukosamin	1,03	1,30
C043	3,30	Thiamin	1,01	1,14
C044	5,41	Xanthin	1,21	1,10
C045	5,11	Karnosin	1,07	1,18
C046	3,31	Guanosin	1,04	1,30
C047	8,37	Akonitát	1,05	1,12
C048	3,57	Cytidin	1,18	0,44
C049	6,67	Cystin	1,11	1,07
C050	4,33	Rafinosa/maltotriososa/melezitosa	0,98	0,89
C051	3,35	Ribitol/arabitol	1,00	1,06
C052	4,24	Guanidinoacetát	1,18	1,17
C053	7,05	N-acetylglutamát	1,15	1,03
C054	3,61	Betain	1,05	1,14
C055	3,45	Spermin	1,15	1,27
C056	2,98	Kreatinin	1,06	1,04
C057	4,08	4-pyridoxát	0,98	1,04

Tab. S1: Pokračování

Označení	t <sub>r</sub> [min]	Analyt	FUT/IUT	NT/IUT
C058	4,98	Pantothenát	1,09	1,10
C059	4,93	Dihydrothymín	1,10	1,16
C060	6,32	2-aminoadipát	1,14	1,19
C061	3,77	Trigonellin	1,10	1,15
C062	4,10	N,N-dimethylglycin/2-aminoisobutyřát	1,05	1,07
C063	4,33	4-guanidinobutanoát	1,11	1,09
C064	7,06	N-acetylaspartát	0,98	1,15
C065	4,84	N-methylhistidin/3-methylhistidin	1,08	1,36
C066	4,45	3-methylhistidin/N-methylhistidin	1,11	1,16
C067	4,45	Hippurát	0,99	1,09
C068	5,57	Threonát	1,11	1,03
C069	3,63	Galaktitol/Manitol	1,05	1,08
C070	4,88	Kynurenát	1,03	1,03
C071	6,40	Xanthosin	1,53	1,42
C072	5,12	5-oxoproline	0,90	1,01
C073	4,48	Pipekolát	0,97	1,05
C074	4,56	Homoarginin	1,06	1,17
C075	3,83	Pseudouridin	1,07	0,99
C076	3,34	7-methylguanin	1,20	1,34
C077	5,85	Cystathionin	1,18	1,17
C078	3,66	Allantoin	1,06	1,30
C079	7,98	Xanthurenát	1,04	1,23
C080	3,66	Acetylkarnitin	1,03	1,17
C081	7,06	2-hydroxyglutarát	1,07	0,99
C082	3,84	N-acetylputrescin	1,09	1,22
C083	3,07	Oktanoylkarnitin	0,96	0,99
C084	3,35	Butyrylkarnitin	1,01	1,11
C085	3,02	Decanoylkarnitin	0,97	1,10
C086	4,42	N,N-Dimethylarginin	1,16	1,27

Tab. S1: Pokračování

Označení	t <sub>r</sub> [min]	Analyt	FUT/IUT	NT/IUT
C087	3,34	FGAr	1,05	1,14
C088	3,43	AICAr	0,89	1,06
C089	6,00	CAI <sub>r</sub>	0,87	0,92
C090	4,32	3-aminoisobutyrát	1,11	1,18
C091	5,43	Vanilmandelát	1,15	1,14
C092	7,06	5-hydroxyindoleacetát	1,21	1,10
C093	4,45	2-OH-isobutyrát/2-OH-butyrát/4-OH-butyrát	1,06	1,15
C094	7,06	Sebakát	1,12	1,15
C095	7,06	Suberát	1,03	1,01
C096	3,03	Decenoylkarnitin	0,92	1,01
C097	3,06	Dekadienylkarnitin	1,10	0,97
C098	3,25	Dodecanoylkarnitin	1,11	1,17
C099	3,14	Dodecenoylkarnitin	1,12	1,14
C100	4,51	1-methylxanthin	1,08	1,18
C101	3,34	Hydroxypropionylkarnitin	1,03	1,19
C102	3,74	Malonylkarnitin/Hydroxybutyrylkarnitin	0,94	1,19
C103	3,46	Butenylkarnitin	1,06	1,14
C104	3,62	Fumarylkarnitin/Hydroxyvalerxylkarnitin	1,05	1,13
C105	3,25	Valerylkarnitin	1,01	1,13
C106	3,31	Tiglylkarnitin	1,02	1,15
C107	3,45	Glutarylkarnitin/Hydroxyhexanoylkarnitin	0,97	1,30
C108	3,18	Hexanoylkarnitin	1,21	1,20
C109	3,22	Hexenoylkarnitin	0,95	0,98
C110	3,31	Adipoylkarnitin/Hydroxyheptanoylkarnitin	0,86	1,07
C111	3,29	Pimelylkarnitin/Hydroxyoktanoylkarnitin	0,98	1,25
C112	3,11	Oktenoylkarnitin	0,97	0,94
C113	3,01	Nonaylkarnitin	1,02	1,09
C114	4,27	2-hydroxyisovalerát	1,06	1,24
C115	4,42	Isobutyrylglycin/Butyrylglycin	1,17	1,24

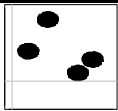
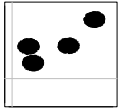
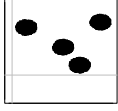
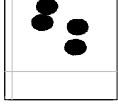

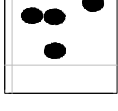
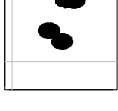
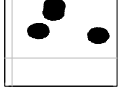
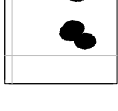
Tab. S1: Pokračování

Označení	t <sub>r</sub> [min]	Analyt	FUT/IUT	NT/IUT
C116	4,14	Hexanoylglycin	1,04	1,32
C117	4,38	Tiglylglycin/3-Methylkrotonylglycin	0,96	1,03
C118	4,29	2-methylbutyrylglycin/Isovalerylglycin	0,92	1,16
C119	4,58	2-Furoylglycin	0,95	0,99
C120	4,46	3-Indolepropionát	1,00	0,97
C121	3,82	Indoxyl sulfát	0,99	0,98
C122	7,22	Succinyladenosin	1,24	1,23
C123	4,05	Myo-inositol	1,13	1,18
		Průměr	1,06	1,13

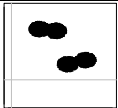
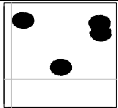

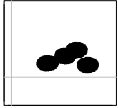
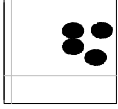
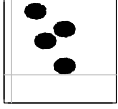

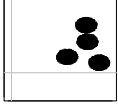
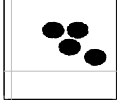
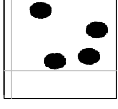
Tab. S2: Průměrné hodnoty ortogonalit jednotlivých typů modelů se 100 body získané různými přístupy výpočtu ortogonalit  $\pm 2SD$  pro 50 modelů stejného typu. Hodnoty jsou uvedeny v %, s výjimkou AN a HN, jejichž výsledkem jsou průměrné nejmenší vzdálenosti mezi nejbližšími sousedy.

Typ modelu	Způsob výpočtu	Min-max	Void-max	WOSEL
	SC	91,7 $\pm$ 8,3	90,8 $\pm$ 9,0	90,2 $\pm$ 8,4
	MA	91,0 $\pm$ 8,8	90,1 $\pm$ 9,2	89,5 $\pm$ 8,6
	CE	77,0 $\pm$ 4,2	77,8 $\pm$ 5,5	78,1 $\pm$ 5,1
	PC	92,9 $\pm$ 10,2	92,9 $\pm$ 10,2	92,9 $\pm$ 10,2
	AE	87,6 $\pm$ 8,3	82,7 $\pm$ 10,0	77,7 $\pm$ 9,1
	CH	75,6 $\pm$ 4,7	70,5 $\pm$ 5,2	66,3 $\pm$ 4,6
	AN	0,0638 $\pm$ 0,0036	0,0616 $\pm$ 0,0035	0,0597 $\pm$ 0,0032
	HN	0,0428 $\pm$ 0,0104	0,0414 $\pm$ 0,0096	0,0401 $\pm$ 0,0094
	SC	91,9 $\pm$ 8,7	38,8 $\pm$ 3,6	13,2 $\pm$ 1,5
	MA	91,4 $\pm$ 8,8	38,6 $\pm$ 3,5	13,1 $\pm$ 1,6
	CE	77,5 $\pm$ 4,4	94,6 $\pm$ 2,8	96,4 $\pm$ 4,0
	PC	94,5 $\pm$ 9,4	94,5 $\pm$ 9,4	94,5 $\pm$ 9,4
	AE	87,8 $\pm$ 9,7	29,7 $\pm$ 4,9	4,9 $\pm$ 0,6
	CH	74,8 $\pm$ 4,4	17,2 $\pm$ 1,4	4,2 $\pm$ 0,3
	AN	0,0634 $\pm$ 0,0036	0,0312 $\pm$ 0,0016	0,0150 $\pm$ 0,0008
	HN	0,0438 $\pm$ 0,0086	0,0215 $\pm$ 0,0040	0,0103 $\pm$ 0,0020
	SC	91,9 $\pm$ 8,7	12,5 $\pm$ 1,2	13,0 $\pm$ 1,6
	MA	91,4 $\pm$ 8,8	12,4 $\pm$ 1,2	12,9 $\pm$ 1,6
	CE	77,5 $\pm$ 4,4	96,2 $\pm$ 3,3	96,3 $\pm$ 3,6
	PC	94,5 $\pm$ 9,4	94,5 $\pm$ 9,4	94,5 $\pm$ 9,4
	AE	87,8 $\pm$ 9,7	5,0 $\pm$ 0,6	4,9 $\pm$ 0,6
	CH	74,8 $\pm$ 4,4	4,2 $\pm$ 0,2	4,2 $\pm$ 0,3
	AN	0,0634 $\pm$ 0,0036	0,0151 $\pm$ 0,0008	0,0150 $\pm$ 0,0008
	HN	0,0438 $\pm$ 0,0086	0,0104 $\pm$ 0,0020	0,0103 $\pm$ 0,0020
	SC	91,9 $\pm$ 8,7	17,5 $\pm$ 2,0	14,5 $\pm$ 2,1
	MA	91,4 $\pm$ 8,8	17,4 $\pm$ 2,0	14,4 $\pm$ 2,1
	CE	77,5 $\pm$ 4,4	96,0 $\pm$ 4,1	94,9 $\pm$ 4,2
	PC	94,5 $\pm$ 9,4	94,5 $\pm$ 9,4	94,5 $\pm$ 9,4
	AE	87,8 $\pm$ 9,7	8,4 $\pm$ 1,0	4,9 $\pm$ 0,5
	CH	74,8 $\pm$ 4,4	6,8 $\pm$ 0,4	4,2 $\pm$ 0,3
	AN	0,0634 $\pm$ 0,0036	0,0191 $\pm$ 0,0010	0,0150 $\pm$ 0,0008
	HN	0,0438 $\pm$ 0,0086	0,0133 $\pm$ 0,0026	0,0103 $\pm$ 0,0020

Tab. S3: Přehled syntetických chromatogramů nízkodimenzionálních modelů typu E a hodnot ortogonalit [%] získaných prostřednictvím jednotlivých deskriptorů.

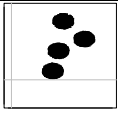
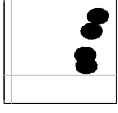

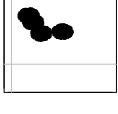
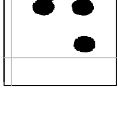
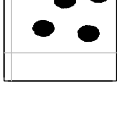
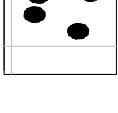
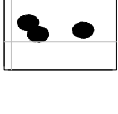
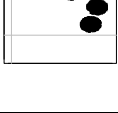

Model	Vzhled	AN	HN	AE	SC	CH	ASCA
1		37,9	23,6	69,8	22,9	44,4	29,5
2		37,3	24,9	49,8	23,4	32,7	29,4
3		37,9	23,9	81,8	23,0	44,8	29,5
4		36,8	25,1	37,7	23,4	25,7	29,2
5		35,8	23,8	46,6	21,5	28,4	27,7
6		37,2	23,9	67,1	22,9	38,3	29,1
7		34,2	21,8	38,0	20,2	22,4	26,3
8		34,9	22,4	48,8	20,3	29,1	26,6
9		33,1	21,3	32,2	19,0	21,0	25,0

Tab. S3: Pokračování

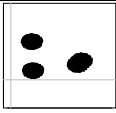
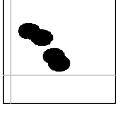
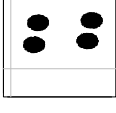
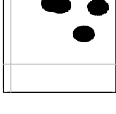

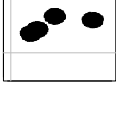
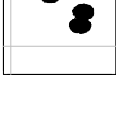
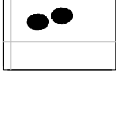

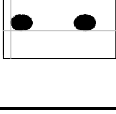
Model	Vzhled	AN	HN	AE	SC	CH	ASCA
10		35,4	22,8	41,2	21,5	24,6	27,7
11		37,5	21,5	86,2	21,7	51,9	28,7
12		33,9	23,7	18,9	20,3	15,9	26,1
13		34,8	23,8	19,8	22,0	16,4	27,6
14		36,3	24,0	28,6	23,1	22,0	28,8
15		37,1	24,6	49,4	23,0	29,3	29,2
16		32,3	20,4	34,0	18,3	20,8	24,3
17		36,6	24,8	34,1	23,3	23,8	29,2
18		36,3	24,8	28,8	22,9	21,5	28,8
19		37,7	22,7	78,4	23,6	46,5	29,8



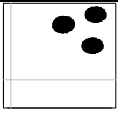

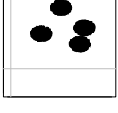
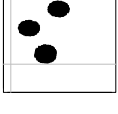
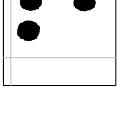
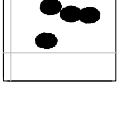
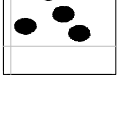
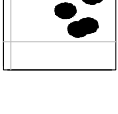
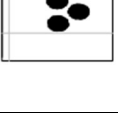
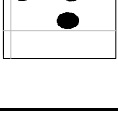
Tab. S3: Pokračování

Model	Vzhled	AN	HN	AE	SC	CH	ASCA
20		36,9	24,2	43,5	23,2	26,6	29,2
21		35,4	23,3	33,3	22,3	18,4	28,1
22		32,9	21,6	23,7	20,0	16,7	25,7
23		33,1	22,3	17,5	20,3	14,4	25,7
24		36,9	23,2	56,5	22,5	33,7	28,8
25		37,8	24,2	63,5	23,6	38,5	29,8
26		37,8	24,0	67,5	23,3	39,7	29,7
27		38,0	23,3	71,7	22,9	40,4	29,5
28		37,2	23,4	45,8	23,3	30,5	29,3
29		38,0	24,1	46,5	22,5	44,4	29,5


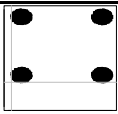
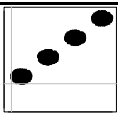
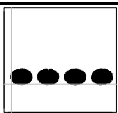
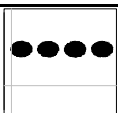
Tab. S3: Pokračování

Model	Vzhled	AN	HN	AE	SC	CH	ASCA
30		34,0	21,2	48,0	19,4	28,1	25,9
31		33,5	22,3	23,9	20,0	15,1	25,9
32		37,7	25,1	56,5	23,4	34,6	29,7
33		34,0	22,1	39,6	20,3	25,0	26,3
34		34,7	21,8	49,1	20,2	26,4	26,5
35		34,0	21,5	36,2	20,1	21,2	26,2
36		35,8	23,2	33,8	22,4	22,3	28,1
37		35,3	19,2	37,1	21,7	23,8	27,7
38		36,8	24,3	50,5	22,6	30,4	29,0
39		36,8	22,1	60,0	21,4	44,8	28,3

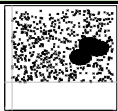
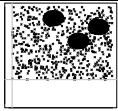



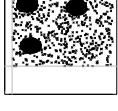
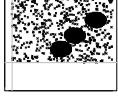
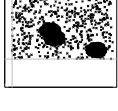

Tab. S3: Pokračování

Model	Vzhled	AN	HN	AE	SC	CH	ASCA
40		32,0	20,6	32,6	18,1	21,7	24,1
41		28,5	18,2	7,9	14,7	8,7	20,5
42		37,2	24,7	42,4	22,7	28,3	29,1
43		32,6	20,9	45,0	18,2	24,8	24,4
44		33,3	20,3	55,6	18,7	30,2	25,1
45		36,5	22,6	39,8	22,5	26,1	28,6
46		37,5	23,7	54,8	22,7	33,2	29,2
47		34,9	22,9	27,9	20,8	20,5	26,9
48		36,6	24,1	36,7	23,0	24,7	29,0
49		38,1	24,8	65,7	23,3	38,4	29,9

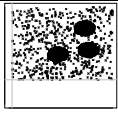
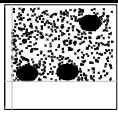
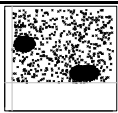
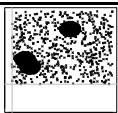

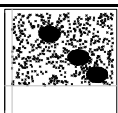
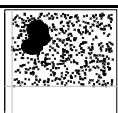
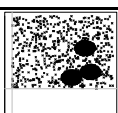
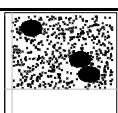
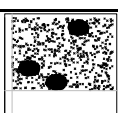
Tab. S3: Pokračování

Model	Vzhled	AN	HN	AE	SC	CH	ASCA
50		33,4	22,0	27,3	19,7	17,4	25,7
51		41,3	24,1	33,0	22,2	98,5	30,3
52		37,4	24,1	27,6	22,7	25,4	29,3
53		36,4	24,0	31,1	22,5	18,8	28,6
54		36,4	23,9	31,1	22,5	18,8	28,5


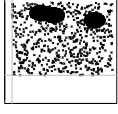
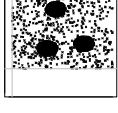
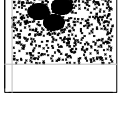
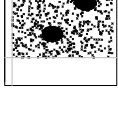
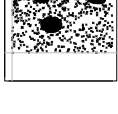
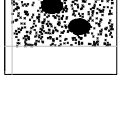

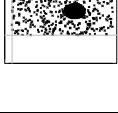

Tab. S4: Přehled syntetických chromatogramů vysokodimenzionálních modelů typu E a hodnot ortogonalit [%] získaných prostřednictvím jednotlivých deskriptorů.

Model	Vzhled	AN	HN	AE	SC	CH	ASCA
1		70,1	25,5	42,1	46,0	96,4	56,8
2		73,1	28,2	58,7	48,5	97,9	59,5
3		69,1	24,5	74,9	45,1	97,9	55,8
4		71,5	26,7	71,8	48,3	97,3	58,8
5		70,7	27,4	64,0	48,1	97,4	58,3
6		71,6	29,1	86,4	48,6	97,7	59,0
7		71,3	29,3	53,5	49,1	96,3	59,2
8		70,8	26,6	58,7	46,7	98,4	57,5
9		71,4	27,2	75,3	49,0	97,3	59,1


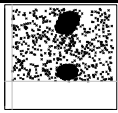
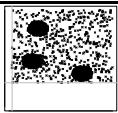

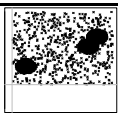

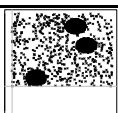
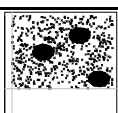
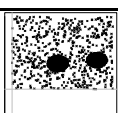

Tab. S4: Pokračování

Model	Vzhled	AN	HN	AE	SC	CH	ASCA
10		71,7	29,8	50,1	49,2	97,2	59,4
11		72,1	28,8	62,9	48,6	97,2	59,2
12		69,5	27,0	66,1	45,7	96,5	56,4
13		71,5	25,6	63,1	47,2	98,4	58,1
14		70,8	25,4	66,3	46,9	96,9	57,6
15		71,9	28,9	60,1	49,7	97,2	59,8
16		69,1	25,6	43,5	46,0	97,8	56,4
17		72,0	28,3	53,0	48,0	97,6	58,8
18		71,9	27,5	61,6	48,8	96,6	59,2
19		70,5	29,7	66,9	47,8	97,0	58,1

Tab. S4: Pokračování


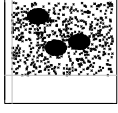
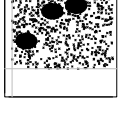
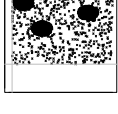

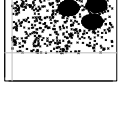


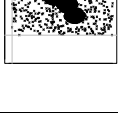

Model	Vzhled	AN	HN	AE	SC	CH	ASCA
20		70,1	27,0	70,8	46,2	96,3	56,9
21		70,8	27,5	63,9	47,5	97,2	58,0
22		71,2	30,4	65,6	48,6	98,2	58,8
23		72,1	28,8	41,4	48,8	95,9	59,3
24		70,8	25,1	61,4	47,1	97,6	57,7
25		71,5	28,4	73,6	47,8	97,6	58,5
26		72,1	27,6	60,6	48,6	97,1	59,2
27		70,8	27,0	45,8	47,4	95,9	57,9
28		72,0	30,4	66,2	49,2	97,6	59,5
29		71,3	28,1	59,1	48,2	97,4	58,6

Tab. S4: Pokračování


Model	Vzhled	AN	HN	AE	SC	CH	ASCA
30		70,4	27,1	44,7	47,7	96,6	57,9
31		70,0	22,1	59,4	46,8	97,3	57,2
32		71,3	25,8	77,0	48,2	98,5	58,6
33		71,7	29,6	51,7	48,8	96,7	59,2
34		71,2	28,8	58,4	48,3	97,1	58,6
35		70,0	25,5	33,7	46,4	97,2	57,0
36		72,2	29,3	66,0	49,1	98,1	59,5
37		71,6	29,2	78,6	48,7	98,2	59,1
38		68,7	23,5	45,9	43,9	98,2	54,9
39		70,7	29,1	56,1	48,6	96,8	58,6



Tab. S4: Pokračování

Model	Vzhled	AN	HN	AE	SC	CH	ASCA
40		72,9	29,7	73,5	49,6	97,8	60,1
41		73,9	30,9	56,7	48,8	98,6	60,1
42		71,9	28,1	66,6	48,4	96,8	59,0
43		71,0	28,5	75,2	47,6	97,3	58,1
44		71,3	29,0	53,5	49,1	98,5	59,2
45		71,2	29,6	48,9	47,9	97,9	58,4
46		70,2	26,6	43,5	47,0	98,1	57,4
47		70,9	30,2	66,5	47,9	96,2	58,3
48		71,7	27,3	37,9	48,3	98,1	58,8
49		69,2	24,8	44,7	45,4	96,9	56,1

Tab. S4: Pokračování

Model	Vzhled	AN	HN	AE	SC	CH	ASCA
50		72,5	28,6	65,1	48,0	97,0	59,0

## 17 Přílohy

Přílohy obsahují 2 reprinty publikovaných článků a text článku v recenzním řízení a profesní životopis autorky.

- Jaroslava Jáčová, Miroslav Jořenek, Kristýna Pospíšková, Lukáš Najdekr, Ludmila Zajoncová, David Friedecký, Tomáš Adam: *Urease-immobilized magnetic microparticles in urine sample preparation for metabolomic analysis by gas chromatography-mass spectrometry*, odesláno do J Chromatogr A
- Jáčová J, Gardlo A, Dimandja JD, Adam T, Friedecký D: *Impact of sample dimensionality on orthogonality metrics in comprehensive two-dimensional separations*. Anal Chim Acta. 2019 Aug 8;1064:138-149. doi: 10.1016/j.aca.2019.03.018. Epub 2019 Mar 12.
- Jáčová J, Gardlo A, Friedecký D, Adam T, Dimandja JD: *Sample-independent approach to normalize two-dimensional data for orthogonality evaluation using whole separation space scaling*. J Chromatogr A. 2017 Aug 18;1511:1-8. doi: 10.1016/j.chroma.2017.06.076. Epub 2017 Jul 3.
- Profesní životopis

# Urease-immobilized magnetic microparticles in urine sample preparation for metabolomic analysis by gas chromatography-mass spectrometry

Jaroslava Jáčová<sup>1,2</sup>, Miroslav Jořenek<sup>3</sup>, Kristýna Pospíšková<sup>4</sup>, Lukáš Najdekr<sup>1,2</sup>, Ludmila Zajoncová<sup>3</sup>, David Friedecký<sup>1,2,\*</sup> and Tomáš Adam<sup>1,2</sup>

<sup>1</sup> Laboratory of Metabolomics, Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacký University Olomouc, Hněvotínská 5, 779 00 Olomouc, Czech Republic

<sup>2</sup> Laboratory for Inherited Metabolic Disorders, Department of Clinical Biochemistry, University Hospital Olomouc, I.P. Pavlova 6, 775 20 Olomouc, Czech Republic

<sup>3</sup> Department of Biochemistry, Faculty of Science, Palacký University Olomouc, Šlechtitelů 27, 783 71 Olomouc, Czech Republic

<sup>4</sup> Regional Center of Advanced Technologies and Materials, Palacký University Olomouc, Šlechtitelů 27, 783 71 Olomouc, Czech Republic

**e-mail:** jaroslava.jacova@gmail.com, MiraCRXJorenek@seznam.cz, kristyna.pospiskova@upol.cz, lukas.najdekr@gmail.com, ludmila.zajoncova@upol.cz, david.friedecky@upol.cz, tomasadam@gmail.com,

## Abstract:

Urea, as an end product of protein metabolism and an abundant polar compound, significantly complicates the metabolomic analysis of urine by GC-MS. We developed a sample preparation method removing urea from urine samples prior the GC-MS analysis. The method based on urease immobilized on magnetic microparticles was compared with the others that are conventionally used (liquid-liquid extraction, free urease protocol), and samples without any treatment. To study the impact of sample preparation approaches on the quality of analytical data, we employed comprehensive metabolomic analysis (using both GC-MS and LC-MS/MS platforms) of standard material based on human urine. Multivariate statistical analysis has shown that immobilized urease treatment provides similar results to a free urease approach. However, significant alterations in the profiles of metabolites were observed in the samples without any treatment and after the extraction. Compared to other approaches that were tested, the immobilization of urease on microparticles reduces both the number of artifacts and the variability of the metabolites (average CV of extraction 19.7%, no treatment 11.4%, free urease

5.0%, and immobilized urease 2.5%). The method that was developed was applied in a GC-MS metabolomic experiment of glutaric aciduria type I, where both known diagnostically important biomarkers and unknowns, as the most discriminating compounds, were found.

**Keywords:** urine sample preparation, immobilized urease, GC-MS, metabolomics

**Highlights:**

- 1) Certified urine is prepared by means of different sample preparation approaches.
- 2) The impact of sample preparation on the urine metabolome is evaluated.
- 3) Urease immobilized on magnetic microparticles is introduced.

## 1 Introduction

The analysis of urine is one of the ways used by physicians to determine the physiological state of a patient. For centuries, urine has assisted in the diagnosing, monitoring, and treatment of a wide range of diseases. It is easy to obtain it non-invasively in large volumes and almost free from interfering proteins and lipids [1]. Compared to other biofluids, metabolites in the urine are usually present in higher levels because of the concentrative effect of the kidneys. High levels of urea in urine interfere in gas chromatography (GC) as separation interference and through the consumption of derivatization agents and therefore compromising the derivatization of analytes [2]. Free acidic hydrogens of analytes resulting from incomplete derivatization interact with the hydroxyl groups of silica columns during the GC separation and cause peak tailing and hence loss of sensitivity and improper quantification.

Many approaches to removing urea prior to analysis have been published. The oldest, liquid-liquid extraction (EX), is still a widely-used method [3]. However, it was primarily intended to isolate selected groups of organic compounds from aqueous phase according to the extraction solvent used (e.g., ethyl ether and/or ethylacetate for organic acids [4,5], hexane for silylated sugars and polyols [6]). EX is mostly performed under acidic conditions (with pH values of 1 to 2) when urea becomes ionized and presents poor solubility in organic phase. Unfortunately, in acidic conditions many urine metabolites turn into an ionic form as well (e.g., creatinine, tryptophan, adenine) and eventually become insoluble (e.g., orotic acid, xanthine, uric acid). Therefore, because of the general absence of many metabolites, EX is not an optimal choice for metabolomic analysis. For the analysis of the complete urine metabolome, another approach utilizing urease has to be used to get rid of urea [7].

The enzyme urease (from *Canavalia ensiformis*, E.C. 3.5.1.5) promises to be an agent for the specific removal of urea from a urine sample. Urease decomposes urea into carbon dioxide and ammonia, which do not interfere with analyses. The first application of urease for the depletion of urea from a urine sample was published in 1964 by Wells, Chin, and Weber [2], who showed that urease pre-treatment eliminates the incomplete silylation of sugars caused by urea. Three decades later, Shoemaker and Elliott [8] applied the urease approach for the automated screening of urine samples for inborn errors of metabolism covering the majority of urine metabolites, namely carbohydrates and organic and amino acids. This method offered good repeatability and accurate reflection of quantitative differences and the expected ratios among different urine volumes. Moreover, qualitatively higher numbers of metabolite identifications were observed [9].

The urease protocols have three main drawbacks. First, most of the protocols using urease treatment are quite laborious [2,10]. Second, urease is added to the sample and, after the depletion of the urea, removed by precipitation [8,9,11,12] causing artifacts (e.g., urease stabilizers or protein residues) to be introduced into the sample [8,13]. Third, the distortion of the concentrations of several metabolites has been reported after urease treatment [9,14]. To simplify and purify the application of urease, urease has been immobilized on various materials, specifically polymers (e.g., chitosan, gelatine) and used as a biosensor for the quantification of urea in aqueous solutions, blood, urine, milk, or alcoholic beverages [15]. The immobilization has also enabled the reuse of the enzyme and has improved the stability of the urease [16,17]. On the basis of previous work [18,19] we suppose that the utilization of magnetic materials as carriers for immobilization further improves manipulation with the enzyme.

The aim of this work was to develop a new protocol for GC-MS metabolomic analysis of urine, where the sample preparation is based on urease immobilized on magnetic microparticles. The immobilized urease protocol was compared to commonly used sample preparations by means of targeted and untargeted metabolomic approaches on a certified urine sample.

## 2 Materials and methods

### 2.1 Chemicals and reagents

N,O bis(trimethylsilyl)trifluoroacetamide with 1% trimethylchlorosilane (BSTFA), O-ethylhydroxylamine hydrochloride, ethyl acetate, pyridine, 4-phenylbutyric acid, mevalonic acid sodium salt, mevalonolactone, N-acetylaspartic acid, an alkane standard mixture for performance tests of GC systems (standard C<sub>10</sub>-C<sub>40</sub>) and LC-MS grade solvents (water, acetic acid, ammonium hydroxide, acetonitrile and methanol) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Chlorotrimethylsilane (TMCS) was shipped by Merck KGaA (Darmstadt, Germany), anhydrous sodium sulphate, sodium chloride, and hydrochloric acid by mikroCHEM (Pezinok, Slovakia), hexanoyl glycine by VU Medical Center (Amsterdam, Netherlands), and the standard material Control Organic Acids (COA) from MCA Laboratory (Winterswijk, Netherlands). All the chemicals used were of analytical grade or higher. A urease immobilized on MG 100 magnetic cellulose microparticles (Iontosorb, Ústí nad Labem, Czech Republic) was developed at the Department of Biochemistry, Faculty of Science, Palacký University, according to the immobilization procedure of Namdeo and Bajpai [20]. A detailed description of the chemicals, methods, and results of the immobilization procedure is provided in the Supplementary Data (Tabs. S1, S2, Figs. S1, S2).

### 2.2 Sample preparation

To evaluate the effect of sample preparation based on immobilized urease on the human urine metabolome, two different materials were analyzed: the standard material Control Organic Acids (COA) from MCA Laboratory and real urine samples. The COA was used for the metabolomic analysis (GC-MS- and LC-MS-based). It COA consists of pooled urine spiked with 23 analytes, mostly organic acids. The COA was reconstituted with 10 mL of water just before use. An amount of COA corresponding to a concentration of creatinine of 1 mmol/L was analyzed by means of GC-MS after four different common sample preparation techniques (acidic extraction into ethyl acetate, free urease treatment, immobilized urease treatment, and no treatment) and by LC-MS/MS. Each experiment was performed in six replicates.

The developed protocol based on immobilized urease was applied to an untargeted GC-MS metabolomic study of the urine of children suffering from glutaric aciduria and healthy controls. Urine samples of the GAI patients and healthy controls were obtained from the Laboratory for Inherited Metabolic Disorders, University Hospital Olomouc, Czech Republic. All the patients and controls gave their informed consent to participation in the study before urine sampling in accordance with the Helsinki Declaration. The samples were split into several aliquots and stored at -80 °C until the analysis. An amount of a sample corresponding to a concentration of creatinine of 1 mmol/L was used for the analysis.

### 2.2.1 Samples that were not subjected to any treatment (NT) for GC-MS

An internal standard (4-phenylbutyric acid, 10 µL, 2.11 mmol/L in methanol; IS) was added to the COA (248 µL), frozen (-80 °C; 30 min) and freeze-dried. The derivatization was performed in two consecutive steps: ethoximation (10 µL of solution in pyridine, 56 mg/mL, and 140 µL of pyridine) for 60 min at 50 °C, silylation (150 µL of silylation mixture BSTFA+TMCS:TMCS, 25:1 (v/v)) for 20 min at 70 °C.

### 2.2.2 Extraction (EX) for GC-MS

The COA (495 µL) was transferred into a standard glass tube containing 20 µL IS. The sample was diluted with water up to 2 mL; O-ethylhydroxylamine hydrochloride (0.03 g) was added to derivatize the ketones and mixed carefully. The samples were incubated for 30 minutes at room temperature. Hydrochloric acid (100 µL; 6 mol/L) and sodium chloride (1.0 g) were added to adjust the pH (1-2) and ionic strength. Liquid-liquid extraction by means of ethyl acetate (4 mL) was performed three times with shaking for 10 min. The upper phase of the solvent was transferred into a new glass tube. In order to remove the water from the samples anhydrous Na<sub>2</sub>SO<sub>4</sub> (1.0 g) was added to each tube, mixed, and left standing for 10 minutes. The solvent phase was decanted into a new glass tube, evaporated to dryness under a nitrogen flow at 30 °C, washed out from the tube in three steps (2x 400 µL of methanol; 300 µL of acetone) using a vortex, and transferred into a GC-MS vial. The sample was evaporated to dryness under a nitrogen flow at 30 °C, then dissolved in a derivatization mixture (BSTFA:pyridine:TMCS, 25:25:1 (v/v/v), 300 µL), capped under a nitrogen atmosphere, and vortexed. All the vials were kept at 70 °C for 20 min.



### 2.2.3 Free urease treatment (FUT) for GC-MS

The reaction mixture, composed of COA (495  $\mu\text{L}$ ), urease (60  $\mu\text{L}$ , 0.6 mg/mL), and water (445  $\mu\text{L}$ ), was stirred in a thermoshaker (37  $^{\circ}\text{C}$ ; 72 g) for 40 min. The reaction mixture (500  $\mu\text{L}$ ) was transferred into a clean microcentrifuge tube and 1 mL of cold methanol (-20  $^{\circ}\text{C}$ ) was added. The sample was centrifuged (14,000 g; 8 min). The supernatant was put into a glass vial containing IS (10  $\mu\text{L}$ ), frozen (-80  $^{\circ}\text{C}$ ; 30 min), and freeze-dried. The derivatization procedure was as same as for the NT samples (Chapter 2.2.1).

### 2.2.4 Immobilized urease treatment (IUT) for GC-MS

The reaction mixture consisted of COA (495  $\mu\text{L}$ ), microparticles (100 mg) with immobilized urease, and water (1505  $\mu\text{L}$ ). The enzyme reaction was performed while the mixture was being stirred in a thermoshaker (37  $^{\circ}\text{C}$ ; 72 g) for 40 min. The magnetic microparticles were separated at the bottom of the vessel using a magnetic separator. The reaction mixture (1 mL) was added to a glass vial with IS (10  $\mu\text{L}$ ), frozen at -80  $^{\circ}\text{C}$  for 30 min, and freeze-dried. The derivatization procedure was the same as for the NT samples (Chapter 2.2.1).

### 2.2.5 Samples for LC-MS analysis

LC-MS analysis of the IUT, FUT, and NT samples was performed. The COA (495  $\mu\text{L}$ ) was mixed with 445  $\mu\text{L}$  of water and 500  $\mu\text{L}$  of the mixture was transferred into a vial and diluted with methanol in the ratio of 1:3 (v/v). No internal standard was added.

## 2.3 Metabolomic analysis

In order to evaluate the differences in the metabolomic data caused by the preparation of the samples, a metabolomic workflow was applied (details shown in the Supplementary Data, Chapter 2: Metabolomic analysis). A comprehensive metabolomic approach consisting of both a GC-MS and an LC-MS/MS platform was used. The statistical analysis involved univariate (e.g., box plots, t-tests, and correlation analysis) and multivariate unsupervised (PCA) and supervised methods (PLS-DA).

### 2.3.1 GC-MS analysis and data processing

Ethoximated trimethylsilylated urine samples (1  $\mu\text{L}$ ) were analyzed with a Pegasus 4 instrument (LECO, St. Joseph, Michigan, USA) equipped with a ZB-5MS column (30 m; 0.25 mm; 0.25  $\mu\text{m}$ ; Phenomenex, Torrance, USA) using a slow linear temperature ramp (5  $^{\circ}\text{C}/\text{min}$  from 80 to 320  $^{\circ}\text{C}$ ). Helium (constant flow 1 mL/min) served as a carrier gas. The data acquisition was performed in the mass range 50-500  $m/z$  at a frequency of 5 Hz. The GC-MS data analysis was performed with ChromaTOF (LECO), XCMS [21], MS Excel 2016 (Microsoft, Redmond, USA), MedCalc (version 16.4, MedCalc Software bvba, Ostend, Belgium), and SIMCA (version 13.0.3, Umetrics, Umea, Sweden). Details are shown in the Supplementary Data.

### 2.3.2 LC-MS analysis and data processing

The COA urine samples (2  $\mu\text{L}$ ) were injected into an UltiMate 3000 RS (Dionex, Sunnyvale, CA, USA) coupled to a triple quadrupole 6500 tandem mass spectrometer (AB Sciex, Foster City, CA, USA), equipped with a Luna  $\text{NH}_2$  column (3.0  $\mu\text{m}$ , 2 x 100 mm; Phenomenex, Torrance, USA; 35  $^{\circ}\text{C}$ ) in normal aqueous phase mode and gradient regime (from 95% ammonium acetate to 95% acetonitrile) at a flow rate at 0.3 mL/min. The previously published method was used [22]. The abundance of 346 targeted metabolites was measured in scheduled multiple reaction monitoring mode with prolonged dwell times. Both quadrupoles were set at unit resolution. The peaks were detected and identified with the MultiQuant 3.0 software (AB Sciex, Foster City, CA, USA), analyzed as compositional data with centered logratio (clr) transformation, and mean centered and processed using a home-made R-package implementing the LOESS (local regression) signal correction method. Details are provided in the Supplementary Data.

## 3 Results and discussion

Urine is a complex material containing thousands of compounds with different chemical properties. There is no universal technique suitable for the analysis of such a wide spectrum of analytes, and therefore two different metabolomic platforms, GC-MS and LC-MS, were applied. Despite the fact that the developed sample preparation method is primarily intended for GC-MS technique, we also applied LC-MS for evaluation of the results as previously published by Kind et al. [14]. The combination

of GC-MS and LC-MS techniques is best suited to cover the changes in urine metabolome by enabling the detection of both lipophilic and hydrophilic metabolites present in the sample.

Since the urea removal is substantial for the GC-MS platform, the stress was put on GC-MS analysis: both targeted and untargeted metabolomic approaches were performed by means of GC-MS analysis. LC-MS/MS analysis showed the possible impact of the sample preparation methods on various metabolites which are not commonly detectable by means of GC-MS analysis. The effect of IUT versus FUT, EX, and NT on the urinary metabolome was evaluated. The certified COA material, which consists of pooled urine spiked with 23 analytes (mostly organic acids), was used. Moreover, since LC-MS on polar stationary phase was also employed, the possible decrease in the matrix effect caused by an abundant peak of urea in the urease-treated samples and the impact on non-GC metabolites was determined. Finally, the new protocol was applied to real urine samples in a metabolomic experiment on glutaric aciduria I patients.

### 3.1 Feasibility of sample preparation protocols

The GC-MS sample preparation protocols that were tested differed in their feasibility. EX required about six working hours. However, the results were available on the same day. The urease protocols took about one hour of handling, but overnight freeze-drying was necessary. The lack of a precipitation step in IUT is an advantage compared to FUT. The time requirements and complexity of the protocol were reduced; the preparation of the IUT sample prior to derivatization took 75% less time than was the case with FUT.

The IUT protocol was developed as an alternative to FUT that would allow the automation of the sample preparation. IUT is easy and straightforward and could be implemented into automated sample preparation stations. The protocol for the preparation of magnetic microparticles is provided in the Supplementary Data.

### 3.2 Retention time reproducibility

The reproducibility of retention times is a basic requirement in terms of the proper alignment and retention index-based identification of metabolites [23]. In order to evaluate the impact of sample treatment, targeted metabolomic analysis was applied.

With GC-MS, all 23 metabolites declared in the certified COA material were found in the samples prepared by means of all the preparation techniques that were tested. Twenty metabolites were identified on the basis of the NIST mass spectral database library (NIST, 2008) with a match greater than 700 and a retention index match with the GOLM database (when possible). The identification of N-acetylaspartic acid (not present in NIST) and hexanoylglycine (tailing peak of low intensity, 9  $\mu\text{mol/L}$ ) was approved via a commercially available standard. Ethylmalonic acid in the NT samples had a low score of identification ( $\sim 300$ ) as a result of its co-elution with a massive peak of urea. It was identified by the presence of specific fragments ( $m/z$  217 and 261) in a given retention time window.

In LC-MS/MS analysis, 123 metabolites in the urine sample prepared by means of IUT, FUT, and NT were detected. An extraction approach was excluded from the experiment because of the necessity of ethoximation prior to the extraction step. Ethoximated derivatives do not fit the MRM transitions of the targeted LC-MS/MS method. On the other hand, the exclusion of ethoximation step significantly changes the extraction efficiency of metabolites.

The retention indices (RI) and retention times, respectively, of the target metabolites were consistent in all the samples analyzed by means of both GC-MS (Tab. S3) and LC-MS/MS (Tab. S4). The average CVs of the RIs were below 0.01% for all the sample preparation techniques. The highest relative differences in retention times between groups were obtained for metabolites inside the retention window of urea of the NT samples, e.g., ethylmalonic acid (0.10%), glyceric acid (0.08%), and mevalonolactone (0.14%). This increase in the RI was caused by the overloading effect of the urea. The RI values differed slightly from the values declared in the GOLM database (Tab. S3), because a slower-temperature GC ramp (4  $^{\circ}\text{C}/\text{min}$  instead of 9  $^{\circ}\text{C}/\text{min}$ ) and ethoximation (instead of methoximation) were applied. For example, for glycolic, sebacic, and 2-ketoglutaric acid RI 1071, 1893, and 1621 were obtained instead of RI 1067, 1888, and 1577 from GOLM (Tab. S3). In the LC-MS/MS analyses, the retention times of the metabolites were consistent, regardless of the sample preparation approach used (Tab. S4). The average coefficient of variation for retention times was 0.7% for all the sample preparations. The maximum and average relative differences vs. medians of all the sample preparations were 0.49% and 0.03%, respectively (for IUT vs NT).

### 3.3 Analytical variability

The sample preparation approaches also differed in the variability of the responses of the target metabolites from the certified COA material. With GC-MS, the median of CV was 5.0, 2.5, 19.7, and 11.4% for FUT, IUT, EX, and NT, respectively (Tab. S5). The CV for FUT was comparable to previously published results (6.1%) for 250 metabolites [9]. Hexanoylglycine was found to be the most variable metabolite (Tab. S5) for all the sample preparation approaches (22-119%) due to chromatography (the low intensity and peak tailing), not the sample preparation. The variability of other metabolites was lowest for IUT. One of the main benefits of FUT, the lowering of variability [9], was not only preserved but even improved by IUT. It was probably caused by simpler sample preparation excluding the freeze-drying of methanol samples (resulting in the formation of bubbles and dispersion of the sample) and by the minimum number of artifacts introduced into the sample (a simpler matrix and thus more efficient deconvolution).

In LC-MS/MS, the variability of the responses (Tab. S6) was comparable for all approaches, with the median ranging from 9.5 to 11.7%. Lactose was found to be the most variable metabolite (66.3%) since as a urease-introduced artifact (see Chapter 3.5) it was present in the NT samples only in trace amounts. The increased %CV values of amino- and thio-compounds in IUT (compared to the FUT and NT samples) could indicate a variable interaction of these functional groups with the magnetic carrier of the urease (e.g., methionine, 2-aminoadipate, panthothenate, carnitines; Tab. S6). However, this phenomenon was not observed in the samples analyzed by means of GC-MS (pyroglutamic acid, tiglylglycine, N-acetylaspartic acid; Tab. S5).

### 3.4 Metabolome coverage

The number of peaks in GC separation and the overall chromatographic pattern differed significantly according to the sample preparation technique used (Figs. 1 and S3). The lowest numbers of chromatographic peaks and altered abundances were found in the EX samples. The highest number of peaks was observed in the FUT samples, as a result of the presence of artifacts at the end of the analytical window (not found in IUT). The rest of the FUT and IUT separation was comparable (Figs. 1C and 1D). The highest TIC abundance was obtained for NT samples (Fig. S3) as a result of their possessing the most complex and concentrated matrix, in which the most prominent peak at a retention time of 12-15 min belonged to urea (Fig. 1C). This huge peak caused a retention time shift of consecutive peaks

(ethylmalonate, glycerate, fumarate, mevalonolactone; Tab. S3) and masked several peaks detected in other separation approaches (Fig. 1D).

In LC-MS/MS analysis the same peak count of 123 was found for IUT, FUT, and NT (EX analysis was not performed, as explained in Chapter 3.2). However, in the FUT samples (both blank and COA), there were increased abundances of sugars: glucose and lactose (see Chapter 3.5).

### 3.5 Artifacts

Artifacts are compounds that are present in the analytical results that do not originate from the sample itself [24]. They could be formed by a derivatization agent (e.g., by-products, unexpected derivatives), introduced by the chromatographic system (e.g., column bleed, products of in-source reactions or reactions in the injector) or during the sample preparation (e.g., residua of reagents, solvent impurities) [14]. Urease is known to introduce artifacts into the analysis [9]. However, the presence of a carrier of immobilized urease may increase their incidence.

In order to compare number of artifacts introduced by the approach, a two different experiments were performed: 1) Comparison of blank samples that revealed 18 peaks of artifacts in FUT, nine in IUT and four in EX ( $S/N > 1000$ ; Fig. 1A,B); 2) An untargeted GC-MS-based metabolomic analysis of COA (Figs. 1C,D; 2C,D; S4-S13) that identified the artifacts (M204T41, M191T42\_1, M191T42\_2, M257T51, M129T38, and M103T42) as the most discriminating compounds between the FUT and IUT samples (Figs. 2D, S6, S7). Abundant peaks at the end of the analytical window (retention time above 37 min, RI above 2430, Fig. 1C), were identified as sugar-like compounds ( $m/z$  204) and probably disaccharides ( $m/z$  361). Since the fragmentation pattern of the same group of saccharides is very similar (e.g., hexoses), it is not possible to identify them via the NIST or GOLM databases. The LC-MS/MS method identified the peak through the SRM transition as a maltose, lactose, sucrose, or melibiose. In previous studies, a significantly increased abundance of lactose, an urease stabilizer, was reported [8,13,25]. Therefore, it can be assumed that there are also multiple peaks of lactose in the FUT samples. These peaks were not detected by GC-MS in the IUT samples, probably as the result of the multiple washing steps enabled by the immobilization of the urease. The more sensitive LC-MS/MS technique identified the presence of lactose in all the sample preparations, even the NT samples. However, the abundance was 57 and 1.5 times higher in FUT and IUT than in NT, respectively (Tab. S7).

Two significant (both of them had score of Variable Importance in Projection of PLS-DA model (VIP) greater than 2 and therefore are significant [26]) non-sugar artifacts were found in FUT (M100T10,

M148T10\_2), one in the IUT samples: M116T10 (Figs. 1A, S6, S7). There was no match either with NIST or with GOLM. The occurrence of  $m/z$  100 suggested the presence of compounds with amino-groups. These artifacts might be residual from the urease (inefficient deproteinization or immobilized urease degradation). It is worth mentioning that when a metabolomic method was used to evaluate the artifacts, statistically significant differences between the sample preparation approaches were obtained. Therefore, the feature with low intensity is also considered important (Fig. 1A, M257T51).

However, it can be concluded that the carrier used for the immobilization of urease did not increase the number of artifacts introduced. By contrast, the incidence of artifacts was lowered by multiple washing of the urease before the sample application.

### 3.6 Alteration of metabolome

The urease protocol is accused of altering the abundances of metabolites [14]. Therefore, this parameter was explored in detail, in three consecutive steps: (1) PCA to visualize a similarity/dissimilarity (grouping trends and outliers in data) of all the sample preparation approaches used, (2) correlation analysis to reveal the overall match between pairs of approaches (e.g., the same trend in all metabolites, lowering of the abundances of several metabolites, completely different patterns), and (3) PLS-DA for detailed data exploration of pairs of approaches (e.g., FUT vs. IUT, FUT vs. NT, FUT vs. EX) to find out the differences in the abundances of a given metabolite. According to the previous part of the work, targeted GC-MS focused on the abundances of 23 metabolites declared in certified material and LC-MS/MS on all 346 metabolites covered 123 metabolites in our targeted metabolomic method. Because IUT represents an alternative to FUT, the greatest emphasis was put on comparing these two approaches. Therefore, the IS was added after the urease treatment, not at the beginning as usual.

PCA (Fig. 2A, 3A) of targeted analyses suggests that extraction into ethylacetate presents a completely different approach that changes the abundances of target analytes (Fig. 2A, PC1 70.0%), namely of glyceric (C07), glycolic (C01), pyroglutamic (C14), and 3-hydroxyisovaleric (C03) acid (Fig. 2B). On the contrary, both urease treatment protocols offer similar results (Fig. 2A, PC2 17.2%; Fig. 3A, PC2 16.6%), and just slightly differ from no treatment samples (Fig. 2A, PC2 17.2%; Fig. 3A, PC1 24.2%), where mainly an altered level of ethylmalonic acid (C06) was observed (Fig. 2B). Indeed, correlation analysis revealed a high degree of similarity in the abundances of metabolites between IUT and FUT ( $r > 0.9871$ ,  $p < 0.0001$ ) and NT and FUT ( $r > 0.9606$ ,  $p < 0.0001$ ) with both targeted GC-MS (Fig. S14) and

LC-MS/MS (Fig. S15). In the comparison of the FUT vs. IUT samples, there was no metabolite outside the line of equality (Fig. S14A), and only one with LC-MS/MS (Fig. S15A). The comparison of the NT and IUT samples (Figs. S14B, S15B) showed an overall decline in the abundances of metabolites in the NT samples with GC-MS, but not LC-MS/MS. EX (Fig. S14C) differed the most from IUT ( $r = 0.5036$ ,  $p = 0.0143$ , degrees of freedom 22). With EX, just eight out of 23 metabolites showed the same abundances as IUT. The abundances of others were affected by the variations in the extraction efficiency into an ethylacetate.

Detailed exploration of the differences between the sample preparation methods was performed via PLS-DA (Figs. S16-S30). In comparison of FUT and IUT samples, 3-hydroxyisobutyric acid was found to be the most discriminating metabolite (increased abundance) for the FUT samples, pyroglutamic acid for the IUT samples with targeted GC-MS analysis (Fig. S18), and lactose (increased in FUT) with LC-MS/MS (Fig. S27). Lactose (C023, increased in FUT), cytidine and glucuronate (C048 and C032, increased in IUT), and uracil (C017, decreased in IUT) discriminate the most between the groups (Fig. 3B). It is worth mentioning that loadings of the metabolites were slightly decentralized in the loading plots of PLS-DA (Figs. S17, S26); therefore, it could be assumed that the overall abundance of the metabolites was slightly lowered for the IUT samples compared to FUT. The average fold change of FUT vs. IUT was 1.05 and 1.06 with GC-MS (Tab. S8) and LC-MS/MS (Tab. S7), respectively. This increase was not uniform because of differences in the solubility of the compounds in methanol (fold change 0.87 of hexoses and 0.82 of riboflavin vs. 1.24 of taurine and 1.53 of kynurenine was obtained). As reported by Chan et al. [12], the methanol serves not only as a deproteinization agent, but also as an extraction solvent. In this work, the hypothesis was supported by the fact that all the hydroxylated target compounds from the COA treated by means of FUT displayed an increase in abundances (Tab. S8). A similar trend was observed for methylated compounds (e.g., methylmalonic and 3-methyl-3-hydroxyglutaric acid). Conversely, FUT caused a decrease in the abundance of pyroglutamic acid [9] which was not so distinct in the IUT samples. It seems that the usage of methanol in the sample preparation of FUT affects the abundance of selected metabolites.

When a comparison of EX and IUT is performed (Figs. S22-S24), an important fact must be considered: EX as a sample preparation was originally developed for the isolation of organic acids and glycine conjugates of interest from urine [27], not for the removal of urea. Therefore, the final composition of the sample is strongly solvent-dependent. If methanol is used, lipids are extracted in substantial amounts [28]. Conversely, if only water is used, sugars and hydroxy acids are completely extracted, mid-polarity compounds are particularly extracted, and lipids remain unextracted in the



pellets [29]. The most commonly used extraction procedure [30], acidic extraction into ethylacetate, was employed in this work. This protocols yields good recoveries (60-107%) for a wide range of clinically significant organic acids and glycine conjugates [31]. However, several compounds, e.g., citric acid [31], methylcitric acid, N-acetylaspartic acid, 2-hydroxybutyric acid, and 3-hydroxybutyric acid [30], suffered from low extraction efficiency. The same phenomenon was observed in this work (Tab. S4), where the abundances of just 12 out of 23 compounds lay in the above-mentioned range from 60% to 107%. The abundance of 2-ketoglutaric acid exceeded this limit. However,  $\alpha$ -keto acids are well known for the fluctuation in their response when they are diluted to more than to creatinine 2 mmol/L [32]. It is worth mentioning that compared to IUT, the urea was not completely depleted by EX; its concentration was just lowered (Fig. 1D). The loading plot of PLS-DA (Fig. S23) shows glyceric, glycolic, and 3-hydroxyisovaleric acid to be the most discriminating metabolites. This finding is in accordance with the above-mentioned work of Tanaka [31], because these compounds belong to group of hydroxy acids similarly to 2-hydroxybutyric, 3-hydroxybutyric, and citric acid. However, according to the course of the VIP plot (Fig. S24), only glyceric acid was a significantly different metabolite.

Analysis of NT samples was biased by enormous abundance of urea. When the comparison of the NT and IUT samples was performed (Figs. S19-S21), ethylmalonic acid (that co-elutes with urea) was the only significantly different metabolite. However, all other metabolites were also compromised in NT samples, as a result of variation in concentration of available derivatization agent (most of it was consumed by urea). Therefore, much greater dispersion of the NT samples on score plot (signaling the versatile derivatization) and decentralization of the data cloud in the loading plot (as a result of low derivatization yield) was obtained.

Compared to the GC-MS, the LC-MS/MS results (Figs. S25, S29) showed the lower statistical significance of PLS-DA models, suggesting smaller differences in levels of analytes between groups. As an example, in the comparison of NT vs. IUT, values of 32.7 vs. 69.8% for LC-MS/MS (Fig. S28) and GC-MS (Fig. S19), respectively, were obtained. Accordingly, the PLS-DA loading (Figs. S26 and S29) and VIP (Figs. S27 and S30) plots represented significantly lower differences with high analytical variability of the metabolites. There was no significant difference between IUT and FUT samples observed (Fig. S24). The comparison of NT and IUT samples revealed three metabolites of VIP > 2: uracil, cytidine, and glucuronate. We did not find the explanation of this phenomenon. However, we have found that it is not caused by the ion suppression, effect of pH on solubility and ionization (as a result of lactam-lactim tautomerism) and cleavage by urease. The average fold change (Tab. S8) showed a

loss of abundance of the metabolites in both urease-treated samples, probably as a result of adsorption on urease (by 13% on average).

### 3.7 Sample complexity

Sample complexity is an important issue in chromatography-based separation techniques. Complex samples require higher separation power of an analytical system [33] and more advanced methods to process the data [23,34]. Matrix effects cause co-elutions, improper deconvolutions, ion suppressions leading to misidentifications or loss of analytes. To date, many sample preparation approaches were developed to decrease the sample matrix. An ideal sample preparation approach influences just the sample matrix, not the analytes. Indeed, in this work, number of peaks (compounds) in final untargeted GC-MS data analysis was the same for all sample preparation approaches tested (199). However, the number of features found in the sample significantly differed, offering the lowest count for EX, the highest for NT samples (Fig. S3).

The lowering of sample complexity via EX was at the expense of alternations of metabolite levels (Fig. 1C,D). Acidic extraction into ethylacetate, beneficial for lowering the urea levels in the sample, did not only affect the urea (Fig. S13, M171T12, obtained RI 1249.3, GOLM RI 1234.6, NIST match 855). A statistically significant decrease in the abundance ( $VIP > 2$ , Fig. S13) was also seen for creatinine (M100T20, obtained RI 1546.6, GOLM RI 1554.40, NIST match 825), glyceric acid (M189T15, obtained RI 1329.4, GOLM RI 1319.9, NIST match 880), unknowns M173T12 (RI 1241.7), and M116T10 (RI 1169.9) as a result of different extraction efficiency in the EX samples. Moreover, the metabolites were also affected by a low pH (1-2 for EX compared to 10 for IUT and FUT), as demonstrated on mevalonic acid/mevalonolactone. It was found that mevalonolactone (M115T16) was preferably formed in acidic conditions, while mevalonic acid (M103T21) in alkaline conditions. When IUT was applied to two samples, healthy urine spiked with either a mevalonolactone or mevalonic acid standard (the final concentration in the sample was 100  $\mu\text{mol/L}$ ), the only peak of mevalonic acid (trimethylsilyl 3,5-bis(trimethylsilyloxy)-3-methylvalerate) was found in both samples (similarity > 900). Mevalonolactone was fully converted (over 98%) into mevalonic acid as a result of the alkaline conditions caused by ammonia. However, the urea 2TMS (M171T12) remained the most discriminating metabolite between EX and urease-treated samples.

On the other hand, the highest sum of features was also not beneficial. The complex sample matrix results to the increase in dTIC within the deconvolution and thus creation of phantom peaks.

The cloud plots (Fig. S3) showed that this event is typical of NT samples, manifested most in the analytical window of urea analytical window (Fig. S3B, RT 10-16 min, RI 1170-1375). In comparison of NT and IUT by untargeted GC-MS, the VIP of PLS-DA (Fig. S10) determined the derivatization products of urea, urea 3TMS (M171T10, obtained RI 1178.0, not present in GOLM, NIST match 769) and urea 2TMS (peak1 M171T12 and peak2 M171T14, obtained RI 1262.7, 1317.0, GOLM RI 1234.6, NIST match 921), as the most discriminating metabolites (VIP > 2). According to the targeted GC-MS, the statistically significant difference was caused by metabolites inside the urea region: ethylmalonic acid (M171T13\_2, obtained RI 1278.1, GOLM RI 1273.3, NIST match 789), phosphate (M189T13, obtained RI 1269.8, GOLM RI 1262.0, NIST match 902), succinate (M172T14, obtained RI 1313.4, GOLM RI 1310.7, NIST match 935), glycerol (M148T13, obtained RI 1274.0, GOLM RI 1262.3, NIST match 938), serine (M116T13, obtained RI 1254.5, GOLM RI 1252.5, NIST match 905), and the unknown M181T13 (RI 1280.7, no match neither in GOLM nor in NIST), which were covered by urea in the NT samples and therefore reduced in them.

Two metabolites were significantly increased in NT: mevalonolactone (M115T16, obtained RI 1364.0, GOLM RI 1378.7, NIST match 905) and uric acid (M441T32, obtained RI 2095.7, GOLM RI 2094.6, NIST match 753). The statistically significant increase in the abundance of mevalonolactone was discussed earlier in this chapter. The lowering of the abundance of uric acid in the FUT and IUT samples was probably caused by the initial addition of water to assure the concentration of creatinine 1 mmol/L and thus reducing of its solubility. In the NT samples, no water was added and an amount of urine corresponding to the concentration of creatinine 1 mmol/L was directly transferred into the GC-MS vial, spiked with IS, and freeze-dried.

It is worth mentioning that Webb-Robertson et al. [9], Shoemaker and Elliot [8], and Matsumoto and Kuhara [11] reported an increase in the abundance of creatinine after FUT, compared to NT samples. This phenomenon was not observed in this work. The comparison of NT vs. IUT with untargeted GC-MS or LC-MS/MS revealed no statistically significant differences in the creatinine level. The fold change was almost the same (an increase by 1% as compared to the average value, metabolite C040, Tab. S7). Shoemaker and Elliot [8] and Matsumoto and Kuhara [11] explained the increase in the abundance of creatinine as a result of the ammonia-caused (high pH) conversion of creatine into creatinine. Since the creatine (26  $\mu\text{mol/L}$ ) against creatinine (2900  $\mu\text{mol/L}$ ) level in the certified COA material was two orders of magnitude lower, the creatine-creatinine conversion could not lead to a statistically significant change in the abundance of creatinine in this case.

The matrix of FUT and IUT was comparable, the artifacts (Fig. 2D, M204T41, M191T42, M257T51, see chapter 3.5), phthalate, and cresol (Fig. 2D, M149T29 and M165T9, respectively), which can be also considered as artifacts, were found to be the most discriminating analytes. The abundances of three metabolites were significantly increased in IUT (Fig. S7): creatinine 3TMS (M100T20; obtained RI 1546.6, GOLM RI 1554.40, NIST match 825, identification approved via the analysis of the corresponding standard), unknown M271T23 (RI 1649.9) and alanine 2TMS (M116T8; obtained RI 1098.8; GOLM RI 1087.4, NIST match 922). The increased abundance of alanine is beneficial, since it could refer to the minor protein contamination of the inlet by IUT. Amino-TMS compounds are most susceptible to loss, as their ratio of formation of nitrogen-silicon bonds and decomposition in the injector very much depend on the total sample matrix [29]. The decrease in the abundance of creatinine in the FUT samples was probably caused by its loss during the precipitation or discrimination of forming N-trimethylsilyl derivatives because of the matrix effect in the inlet. The difference in the RI that was obtained and the GOLM-referred RI was probably caused by a pH-evoked tautomeric change in the structure of the creatinine and thus dissimilar chromatographic behavior. This compound is well known for tautomeric changes in dependence on the pH or type of solvent [35].

### 3.8 Application of immobilized urease to metabolomics of glutaric aciduria patients

In order to evaluate the IUT method, a metabolomic study of patients suffering from glutaric aciduria type I (GAI) was performed. Urine samples (n=10) from patients and healthy controls (n=10) were prepared, analyzed, and statistically processed according to the usual metabolomic workflow (details are provided in the Supplementary Data, Chapter 2). A total of 221 compounds was found in the samples. The score plot of PCA (Fig. 4) shows a clear segregation of the GAI patients from the healthy controls. The higher dispersion of the GAI samples could be associated with the clinical state of the diseased patients. The model probability of PCA for the first two principal components was 34.6%. Both PCA and PLS-DA (Figs. 4, S31, S32, and S33) showed the most discriminating significant metabolites (VIP>2) as follows: increased levels of 3-hydroxyglutaric acid (M185T23, obtained RI 1574.8, library match 968), tartaric acid (M102T24, obtained RI 1636.1, NIST match 917), glutaric acid (M261T18, obtained RI 1403.3, GOLM RI 1401.4, NIST match 880) and unknown M111T13 (RI 1213.4), decreased levels of the unknowns M210T14 (RI 1240.0), M224T17 (RI 1361.5), M210T16 (RI 1317.3), M210T17 (RI 1350.3), and M111T21 (RI 1506.9). 3-hydroxyglutaric acid and glutaric acid are well-known

biomarkers of GAI. Tartaric acid is contained in the dietary treatment of GAI patients. The unknowns M210T14, M224T17, M210T16, M210T17, and M111T21 are common components of healthy urine. Neither the GOLM nor the NIST database identified them. This will be the topic of upcoming research.

## 4 Conclusions

Urease immobilized on magnetic microparticles is a promising new approach to the preparation of samples for the GC-MS analysis of urine. It is easy to perform; the sample is prepared in 40 min and freeze-dried. Compared to other traditional sample preparation approaches (free urease treatment, no treatment, and extraction into an ethyl acetate) it provides the lowest analytical variability. For 23 target metabolites of a standard reference material an average CV of 2.5% is obtained.

Immobilized urease offers similar results to free urease. The PCA, correlation analysis, and PLS-DA of the samples analyzed with GC-MS (both targeted and untargeted) and targeted LC-MS/MS do not reveal any significant alteration of the abundances of metabolites, with the exception of the most abundant artifact, lactose, introduced by free urease. In total, a lower number of artifacts (3 vs. 10) is introduced by immobilized urease. However, the presence of artifacts does not cause any retention time shift and has no impact on the identification of metabolites.

On the other hand, the extraction protocol does not remove the total urea load and alters the abundances of metabolites (namely creatinine and glyceric acid) in a standard material based on human urine. In the samples that were not subjected to any treatment, the presence of a huge peak of urea results in column overload, retention time shift, and disabling the accurate identification and determination of metabolites eluting in the retention window of urea.

The application of the method that was developed is demonstrated through the untargeted GC-MS-based metabolomic profiling of urine samples from patients suffering from glutaric aciduria type I. Immobilized urease provides an advantageous alternative to traditional sample preparation techniques.

### Acknowledgements

This work was supported by the Czech Science Foundation Grant [18-12204S]. The infrastructural part (Institute of Molecular and Translational Medicine) was supported by NPU I (LO1304).

## References

- [1] S. Bouatra, F. Aziat, R. Mandal, A.C. Guo, M.R. Wilson, C. Knox, T.C. Bjorndahl, R. Krishnamurthy, F. Saleem, P. Liu, Z.T. Dame, J. Poelzer, J. Huynh, F.S. Yallou, N. Psychogios, E. Dong, R. Bogumil, C. Roehring, D.S. Wishart, The Human Urine Metabolome, *PLoS One*. 8 (2013). doi:10.1371/journal.pone.0073076.
- [2] W.W. Wells, T. Chin, B. Weber, Quantitative analysis of serum and urine sugars by gas chromatography, *Clin. Chim. Acta*. 10 (1964) 352–359. doi:10.1016/0009-8981(64)90066-X.
- [3] J.D. Shoemaker, W.H. Elliott, Automated screening of urine samples for carbohydrates, organic and amino acids after treatment with urease, *J. Chromatogr. B Biomed. Sci. Appl.* 562 (1991) 125–138. doi:10.1016/0378-4347(91)80571-S.
- [4] S.I. Goodman, S.P. Markey, Diagnosis of organic acidemias by gas chromatography--mass spectrometry., *Lab. Res. Methods Biol. Med.* 6 (1981) 1–158.
- [5] L. Chalmers RA, Lawson AMChapman and Hall, Organic Acids in Man: Analytical Chemistry, Biochemistry and Diagnosis of the Organic Acidurias, *Anal. Biochem.* 127 (1982) 458–460.
- [6] G. Jansen, F.A. Muskiet, H. Schierbeek, R. Berger, W. van der Slik, Capillary gas chromatographic profiling of urinary, plasma and erythrocyte sugars and polyols as their trimethylsilyl derivatives, preceded by a simple and rapid prepurification method., *Clin. Chim. Acta.* 157 (1986) 277–93. doi:10.1016/0009-8981(86)90303-7.
- [7] T. Kuhara, Diagnosis of inborn errors of metabolism using filter paper urine, urease treatment, isotope dilution and gas chromatography-mass spectrometry, *J Chromatogr B Biomed Sci Appl.* 758 (2001) 3–25.
- [8] J.D. Shoemaker, W.H. Elliott, Automated screening of urine samples for carbohydrates, organic and amino acids after treatment with urease., *J. Chromatogr.* 562 (1991) 125–38. <http://www.ncbi.nlm.nih.gov/pubmed/2026685>.
- [9] B.-J. Webb-Robertson, Y.-M. Kim, E.M. Zink, K.A. Hallaian, Q. Zhang, R. Madupu, K.M. Waters, T.O. Metz, A Statistical Analysis of the Effects of Urease Pre-treatment on the Measurement of the Urinary Metabolome by Gas Chromatography-Mass Spectrometry., *Metabolomics*. 10 (2014) 897–908. doi:10.1007/s11306-014-0642-1.
- [10] J.D. Shoemaker, One-step Metabolomics: Carbohydrates, Organic and Amino Acids Quantified in a Single Procedure, *J. Vis. Exp.* (2010) 12–14. doi:10.3791/2014.
- [11] I. Matsumoto, T. Kuhara, A new chemical diagnostic method for inborn errors of metabolism by

- mass spectrometry—rapid, practical, and simultaneous urinary metabolites analysis, *Mass Spectrom. Rev.* 15 (1996) 43–57. doi:10.1002/(SICI)1098-2787(1996)15:1<43::AID-MAS3>3.0.CO;2-B.
- [12] E.C.Y. Chan, K.K. Pasikanti, J.K. Nicholson, Global urinary metabolic profiling procedures using gas chromatography–mass spectrometry, *Nat. Protoc.* 6 (2011) 1483–1499. doi:10.1038/nprot.2011.375.
- [13] R.S. Clements, W.R. Starnes, An improved method for the determination of urinary myoinositol by gas-liquid chromatography, *Biochem. Med.* 12 (1975) 200–204. doi:10.1016/0006-2944(75)90112-X.
- [14] T. Kind, V. Tolstikov, O. Fiehn, R.H. Weiss, A comprehensive urinary metabolomic approach for identifying kidney cancer, *Anal. Biochem.* 363 (2007) 185–195. doi:10.1016/j.ab.2007.01.028.
- [15] B. Krajewska, Ureases. II. Properties and their customizing by enzyme immobilizations: A review, *J. Mol. Catal. B Enzym.* 59 (2009) 22–40. doi:10.1016/j.molcatb.2009.01.004.
- [16] L. Cao, *Carrier-bound Immobilized Enzymes*, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, FRG, 2005. doi:10.1002/3527607668.
- [17] B. Sahoo, S.K. Sahu, P. Pramanik, A novel method for the immobilization of urease on phosphonate grafted iron oxide nanoparticle, *J. Mol. Catal. B Enzym.* 69 (2011) 95–102. doi:10.1016/j.molcatb.2011.01.001.
- [18] A.R. Varlan, W. Sansen, A. Van Loey, M. Hendrickx, Covalent enzyme immobilization on paramagnetic polyacrolein beads, *Biosens. Bioelectron.* 11 (1996) 443–447. doi:10.1016/0956-5663(96)82740-0.
- [19] R.P. Pogorilyi, I. V. Melnyk, Y.L. Zub, G.A. Seisenbaeva, V.G. Kessler, Immobilization of urease on magnetic nanoparticles coated by polysiloxane layers bearing thiol- or thiol- and alkyl-functions, *J. Mater. Chem. B.* 2 (2014) 2694–2702. doi:10.1039/C4TB00018H.
- [20] M. Namdeo, S.K. Bajpai, Immobilization of  $\alpha$ -amylase onto cellulose-coated magnetite (CCM) nanoparticles and preliminary starch degradation study, *J. Mol. Catal. B Enzym.* 59 (2009) 134–139. doi:10.1016/j.molcatb.2009.02.005.
- [21] R. Tautenhahn, G.J. Patti, D. Rinehart, G. Siuzdak, XCMS Online: a web-based platform to process untargeted metabolomic data., *Anal. Chem.* 84 (2012) 5035–5039. doi:10.1021/ac300698c.
- [22] R. Karlíková, J. Šíroková, D. Friedecký, E. Faber, M. Hrdá, K. Mičová, I. Fikarová, A. Gardlo, H. Janečková, I. Vrobel, T. Adam, Metabolite Profiling of the Plasma and Leukocytes of Chronic Myeloid Leukemia Patients, *J. Proteome Res.* 15 (2016) 3158–3166.

- doi:10.1021/acs.jproteome.6b00356.
- [23] Y. Koh, K.K. Pasikanti, C.W. Yap, E.C.Y. Chan, Comparative evaluation of software for retention time alignment of gas chromatography/time-of-flight mass spectrometry-based metabonomic data, *J. Chromatogr. A*. 1217 (2010) 8308–8316. doi:10.1016/j.chroma.2010.10.101.
- [24] J.L. Little, Artifacts in trimethylsilyl derivatization reactions and ways to avoid them, *J. Chromatogr. A*. 844 (1999) 1–22. doi:10.1016/S0021-9673(99)00267-8.
- [25] F. Palmas, M. Mussap, C. Fattuoni, Urine metabolome analysis by gas chromatography–mass spectrometry (GC–MS): Standardization and optimization of protocols for urea removal and short-term sample storage, *Clin. Chim. Acta*. 485 (2018) 236–242. doi:10.1016/j.cca.2018.07.006.
- [26] M. Pérez-Enciso, M. Tenenhaus, Prediction of clinical outcome with microarray data: a partial least squares discriminant analysis (PLS-DA) approach., *Hum. Genet.* 112 (2003) 581–92. doi:10.1007/s00439-003-0921-9.
- [27] D.C. Lehotay, J.T.R. Clarke, P. Renaldo, Organic Acidurias and Related Abnormalities, *Crit. Rev. Clin. Lab. Sci.* 32 (1995) 377–429. doi:10.3109/10408369509084689.
- [28] O. Fiehn, G. Wohlgemuth, M. Scholz, T. Kind, D.Y. Lee, Y. Lu, S. Moon, B. Nikolau, Quality control for plant metabolomics: reporting MSI-compliant studies, *Plant J.* 53 (2008) 691–704. doi:10.1111/j.1365-313X.2007.03387.x.
- [29] O. Fiehn, Metabolomics by Gas Chromatography-Mass Spectrometry: Combined Targeted and Untargeted Profiling, in: *Curr. Protoc. Mol. Biol.*, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2016: pp. 30.4.1-30.4.32. doi:10.1002/0471142727.mb3004s114.
- [30] A. Liu, M.M. Kushnir, W.L. Roberts, M. Pasquali, Solid phase extraction procedure for urinary organic acid analysis by gas chromatography mass spectrometry., *J. Chromatogr. B. Analyt. Technol. Biomed. Life Sci.* 806 (2004) 283–7. doi:10.1016/j.jchromb.2004.03.048.
- [31] K. Tanaka, A. West-Dull, D.G. Hine, T.B. Lynn, T. Lowe, Gas-chromatographic method of analysis for urinary organic acids. II. Description of the procedure, and its application to diagnosis of patients with organic acidurias., *Clin. Chem.* 26 (1980) 1847–53. <http://www.ncbi.nlm.nih.gov/pubmed/7438430>.
- [32] P. Duez, A. Kumps, Y. Mardens, GC-MS profiling of urinary organic acids evaluated as a quantitative method., *Clin. Chem.* 42 (1996) 1609–15. <http://www.ncbi.nlm.nih.gov/pubmed/8855143>.
- [33] J.C. Giddings, Sample dimensionality: a predictor of order-disorder in component peak



- distribution in multidimensional separation., *J. Chromatogr. A.* 703 (1995) 3–15. doi:10.1016/0021-9673(95)00249-M.
- [34] W.B. Dunn, I.D. Wilson, A.W. Nicholls, D. Broadhurst, The importance of experimental design and QC samples in large-scale and MS-driven untargeted metabolomic studies of humans, *Bioanalysis.* 4 (2012) 2249–2264. doi:10.4155/bio.12.204.
- [35] J. Gao, Y. Hu, S. Li, Y. Zhang, X. Chen, Tautomeric equilibrium of creatinine and creatinium cation in aqueous solutions explored by Raman spectroscopy and density functional theory calculations, *Chem. Phys.* 410 (2013) 81–89. doi:10.1016/j.chemphys.2012.11.002.

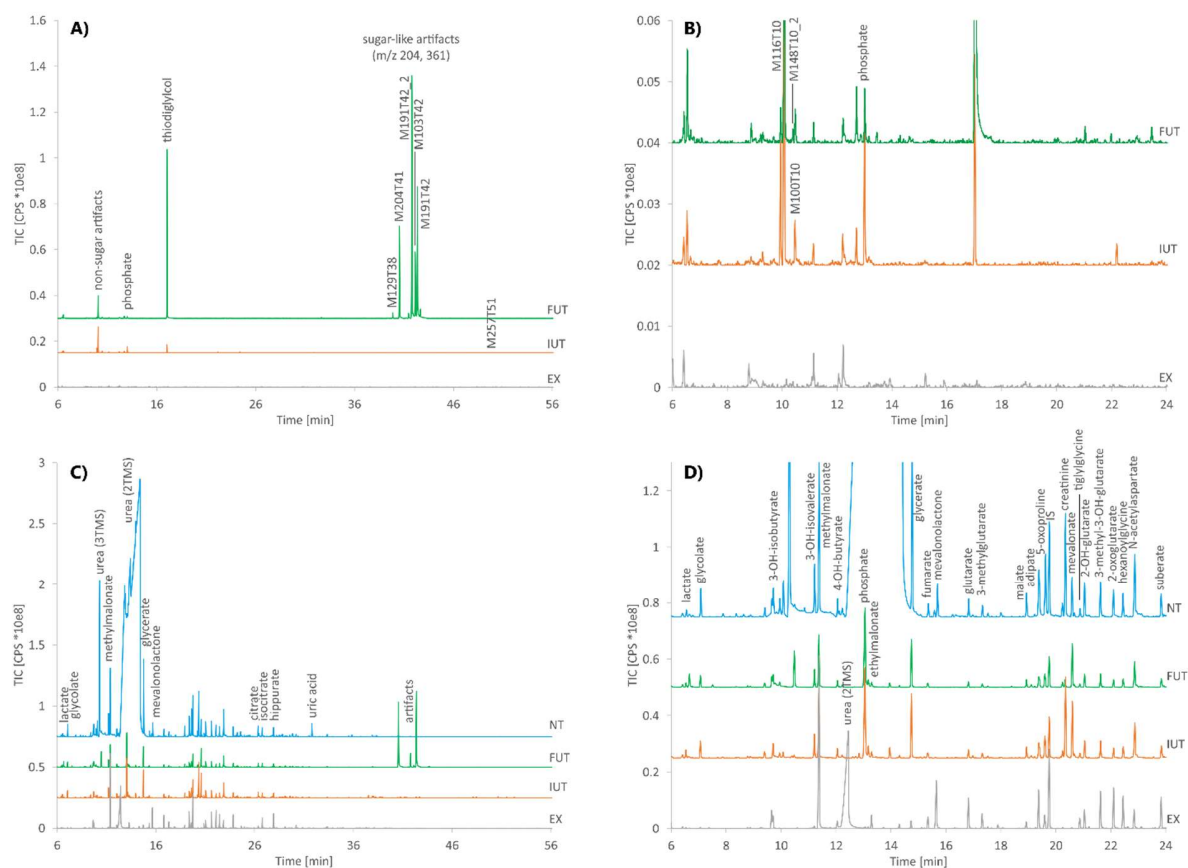


Fig. 1: Comparison of GC-MS analysis of blank samples (A – overall pattern, B – detail) and standard material Control Organic Acids (C – overall pattern, D – detail) from MCA Laboratory (ERNDIM) with different sample preparation techniques: FUT – free urease treatment, IUT – immobilized urease treatment, EX – extraction, NT – samples without any treatment. The standard material is based on human urine, spiked with known levels of 23 metabolites, mostly organic acids.

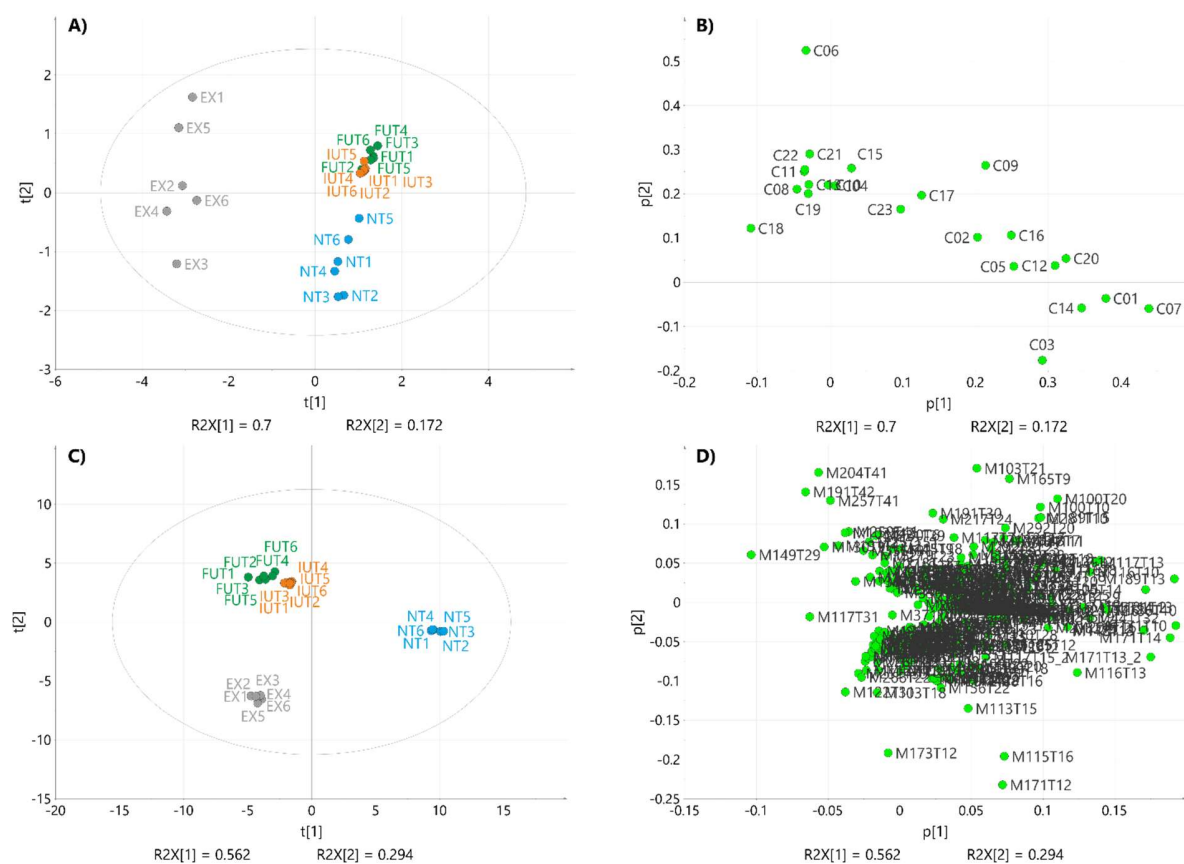


Fig. 2: Score (A, C) and loading (B, D) plots of the first two principal components of PCA of targeted (A, B) and untargeted (C, D) GC-MS analysis of COA treated by means of different sample preparations. Fig. 2B: C07 – glyceric, C01 – glycolic acid, C14 – pyroglutamic acid, C03 – 3-hydroxyisovaleric acid, C06 – ethylmalonic acid; the complete list of the names of the metabolites is shown in Tab. S8. Fig. 2D: M171T12 – urea; M115T16 – mevalonolactone; M103T21 – mevalonic acid; M100T20 – creatinine; M204T41, M191T42, M257T41, M100T10 – artifacts; M149T29 – phthalic acid, M165T9 – cresol.

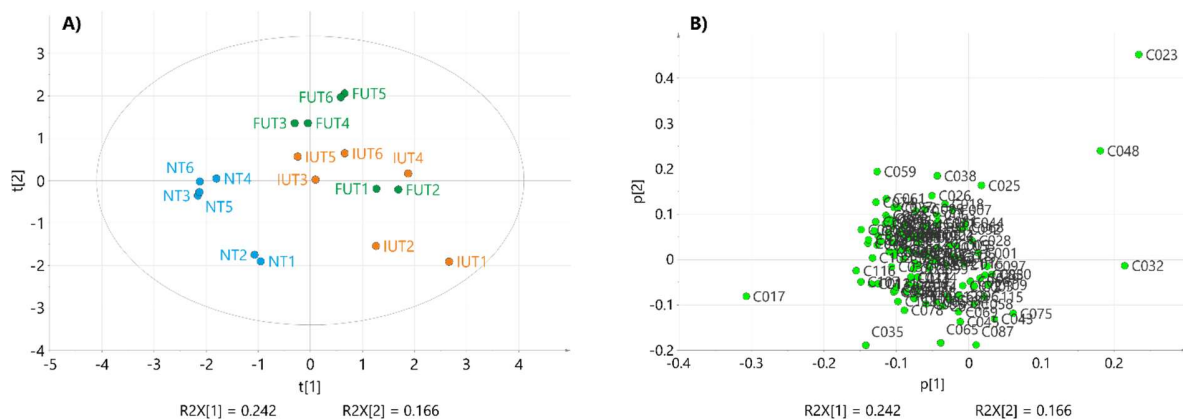


Fig. 3: Score plot (A) and loading plot (B) of the first two principal components of PCA of targeted LC-MS/MS analysis of COA treated by means of different sample preparations. C023 – lactose, C048 – cytidine, C032 – glucuronate, C017 – uracil. The complete list of the names of the metabolites is shown in Tab. S7.

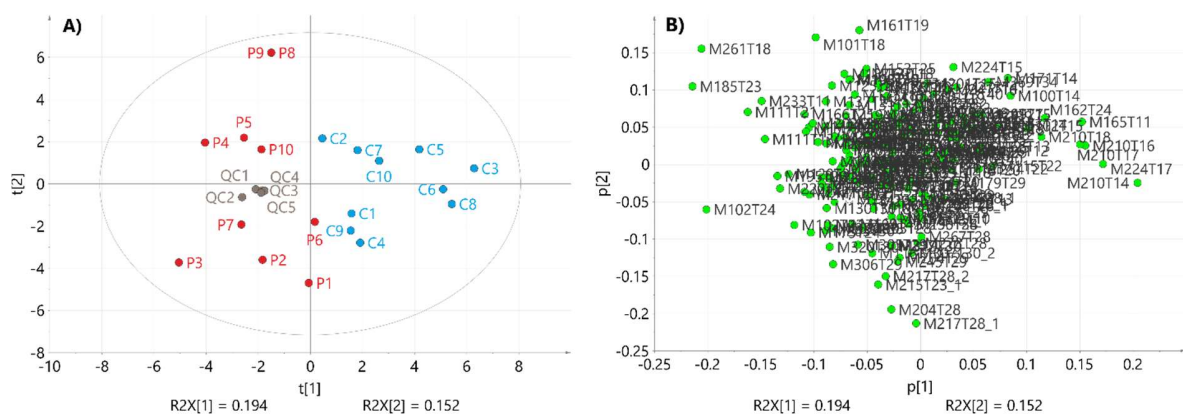


Fig. 4: Score plot and loading plot of the first two principal components of PCA of patients suffering from glutaric aciduria (P1-P10), healthy controls (C1-C10) and quality control samples (QC1-QC5). The biomarkers of GAI: 3-hydroxyglutaric acid (M185T23), tartaric acid (M102T24), and glutaric acid (M261T18).



# Sample-independent approach to normalize two-dimensional data for orthogonality evaluation using whole separation space scaling<sup>☆</sup>

Jaroslava Jáčová<sup>a,b</sup>, Alžběta Gardlo<sup>a,b</sup>, David Friedecký<sup>a,b,\*</sup>, Tomáš Adam<sup>a,b</sup>, Jean-Marie D. Dimandja<sup>c</sup>

<sup>a</sup> Laboratory of Metabolomics, Institute of Molecular and Translational Medicine, Faculty of Medicine and Dentistry, Palacký University Olomouc, Hněvotínská 5, 779 00 Olomouc, Czech Republic

<sup>b</sup> Department of Clinical Chemistry, University Hospital Olomouc, I. P. Pavlova 6, 779 00 Olomouc, Czech Republic

<sup>c</sup> School of Mechanical Engineering, Georgia Institute of Technology, 801 Ferst Dr. NW, 30313 Atlanta, United States

## ARTICLE INFO

### Article history:

Received 31 October 2016

Received in revised form 29 June 2017

Accepted 30 June 2017

Available online 3 July 2017

### Keywords:

Orthogonality

Normalization

Scaling

Separation efficacy

Sample-independency

2D separations

## ABSTRACT

Orthogonality is a key parameter that is used to evaluate the separation power of chromatography-based two-dimensional systems. It is necessary to scale the separation data before the assessment of the orthogonality. Current scaling approaches are sample-dependent, and the extent of the retention space that is converted into a normalized retention space is set according to the retention times of the first and last analytes contained in a unique sample to elute. The presence or absence of a highly retained analyte in a sample can thus significantly influence the amount of information (in terms of the total amount of separation space) contained in the normalized retention space considered for the calculation of the orthogonality. We propose a Whole Separation Space Scaling (WOSEL) approach that accounts for the whole separation space delineated by the analytical method, and not the sample. This approach enables an orthogonality-based evaluation of the efficiency of the analytical system that is independent of the sample selected. The WOSEL method was compared to two currently used orthogonality approaches through the evaluation of *in silico*-generated chromatograms and real separations of human biofluids and petroleum samples. WOSEL exhibits sample-to-sample stability values of 3.8% on real samples, compared to 7.0% and 10.1% for the two other methods, respectively. Using real analyses, we also demonstrate that some previously developed approaches can provide misleading conclusions on the overall orthogonality of a two-dimensional chromatographic system.

© 2017 Elsevier B.V. All rights reserved.

## 1. Introduction

In the last decade, increasing requirements for the analysis of complex samples have led to the development of “comprehensive” multidimensional analytical techniques that are based on the coupling of two or more separation dimensions [1]. The separation mechanisms of these coupled dimensions should be optimized to have significantly different and independent separation selectivities with respect to the target analytes in a sample. The concept that best describes this selectivity independence of the two-dimensional (2D) chromatographic system is termed

separation orthogonality [2,3]. No consensus on a formal definition of an orthogonality metric in chromatography has been reached yet [4]. In this work, orthogonality is considered as a value that describes the degree of complementarity of the separation system towards the sample with respect to the purpose of the analysis.

In chromatography-based 2D separations the range of first and second dimensions differs significantly in magnitude [5–7], thus requiring a scaling of the data. Scaling (which is also called normalization) is a common initial data treatment procedure in the evaluation of the orthogonality in a 2D separation space. It is the mathematical operation that gives the same strength to the first and second separation dimensions. The result of the scaling process is the generation of a uniform 2D retention space, regardless of absolute values (normalized separation space), and also the removal of void or unused spaces in the 2D separation plot, where no peaks are able to elute [8].

The range scaling transformation was described for the first time by Steuer et al. [5]. It converted multisource retention data ( $Rt_1$ )

<sup>☆</sup> Selected paper from the 40th International Symposium on Capillary Chromatography and 13th GC×GC Symposium (RIVA 2016), 29 May–3 June 2016, Riva del Garda, Italy.

\* Corresponding author at: Department of Clinical Chemistry, University Hospital Olomouc, I. P. Pavlova 6, 779 00 Olomouc, Czech Republic.

E-mail address: [david.friedecky@gmail.com](mailto:david.friedecky@gmail.com) (D. Friedecký).

into the retention factors  $X_a$  by the subtraction of the retention time of the unretained component ( $Rt_0$ ), with subsequent division by the analysis time (the retention time of the longest-eluting component ( $Rt_{\max}$ ) minus  $Rt_0$ , Eq. (1)).

$$X_a = (Rt_1 - Rt_0)/(Rt_{\max} - Rt_0) \quad (1)$$

This equation, called normalization according to the retention mean [9], was used in much of the early separation orthogonality work. By 2005, the retention time of the unretained component in Eq. (1) had been replaced by the retention time of the first component to elute ( $Rt_{\min}$ , Eq. (2); [8]). The main benefits of the use of  $Rt_{\min}$  were the simplification of the determination (the void time ( $Rt_0$ ) changes with the temperature in the GC oven) and the sample-specificity (it enables the evaluation of the orthogonality targeted at the given sample).

$$X_a = (Rt_1 - Rt_{\min})/(Rt_{\max} - Rt_{\min}) \quad (2)$$

In an ideal 2D chromatogram, the whole separation space should be uniformly covered by data points [10], which would result in a maximum calculated orthogonality value of 100%. A number of orthogonality evaluation approaches have been published to date [4,10,11], and can be classified into three categories. The first category of approaches, which includes methods such as correlation coefficients [9] and conditional entropy [11], determines the degree of similarity of information brought about by the two coupled dimensions.

The second category of approaches evaluates the coverage of the separation space, and can be further split into two sub-categories: discretized and non-discretized [4]. The discretized methods compartmentalize the separation space into bins. The orthogonality value (%) reflects the ratio of occupied bins to the total sum of bins, divided by the probability of bin matching [8,12]. Non-discretized methods do not require a partitioning of the separation space [13–16] but they differ in their respective operating principles. The nearest neighbor approach [13] computes the average distance to a nearest neighbor for all the peaks present in the separation. The convex hull approach compares the area covered by the peaks to the usable retention area [14,15]. The asterisk equations approach [16] expresses the orthogonality as a combination of the data distribution along four axes crossing the separation space.

Finally, in the third group of approaches the information similarity and separation space usage parameters are combined into one %O value (e.g. the modeling approach [17] or Geometric Approach to Factor Analysis [9]).

The purpose of orthogonality calculations is to determine the degree to which a given chromatographic method is able to resolve a particular sample, which is typically useful in column selection investigations. A target list of analytes or the sample as a whole [18] is used for evaluation of the orthogonality, depending on the purpose of the analysis (e.g. a targeted or untargeted approach in metabolomics). In the event of a sample that is variable in content (many environmental and biological samples) there is a need to develop a more general scaling method. Current orthogonality metrics (involving Min-Max or Void-Max scaling) are only valid for samples with stable compositions.

The aim of this work is to develop a new scaling approach that can be used with virtually every existing orthogonality determination method. This approach, which is termed **Whole SEparation Space ScaLing (WOSEL)**, is intended to provide a less sample-dependent orthogonality estimation. WOSEL is compared with two general scaling approaches, and its impact on orthogonality calculations is evaluated on theoretical models as well as on six different GC  $\times$  GC real analyses of urine, plasma and petrol.

## 2. Materials and methods

Artificial chromatograms were generated *in silico* using the R software (v. 3.3.0, <https://www.r-project.org>). Two models consisting of 100 and 2500 data points were generated in a 2D separation space (89–3000 s  $\times$  1–4 s) to form one oval cluster. A uniform distribution on an ellipse-shaped contour was used for all the data generation. The data was generated into a normalized retention space in order to form a filled circle of appropriate radius (unit-length or quarter-length) with the minimum distance of 0.001 between the data points. The retention coordinates were then recalculated into a real separation space. Since the scales of the first and second dimensions differ in magnitude, oval clusters were created. Six different scenarios were generated (Figs. 1, S3): a single one-unit length cluster (the length of the major axis was equal to the analysis time in the first dimension, the length of the minor axis to the analysis time in the second dimension) and five clusters of the size of one quarter of the unit length (the length of the major axis was equal to the analysis of one quarter of the time in the first dimension, the length of the minor axis to one quarter of the analysis time in the second dimension) differing in terms of their location in the separation space. Each scenario was produced 50 times *via* the same algorithm based on random sampling to evaluate the validity of the model (using standard deviation values).

The retention coordinates of the data points were scaled in three different ways (Fig. 2): Min-Max scaling (using the minimal and maximal retention time obtained in the given dimension, Eq. (2)), Void-Max scaling (using the void time and maximal retention time obtained in the given dimension, Eq. (1)) and the new type of scaling that we call WOSEL (delineated by the void time and end of the method, Eq. (3), see below for further discussion).

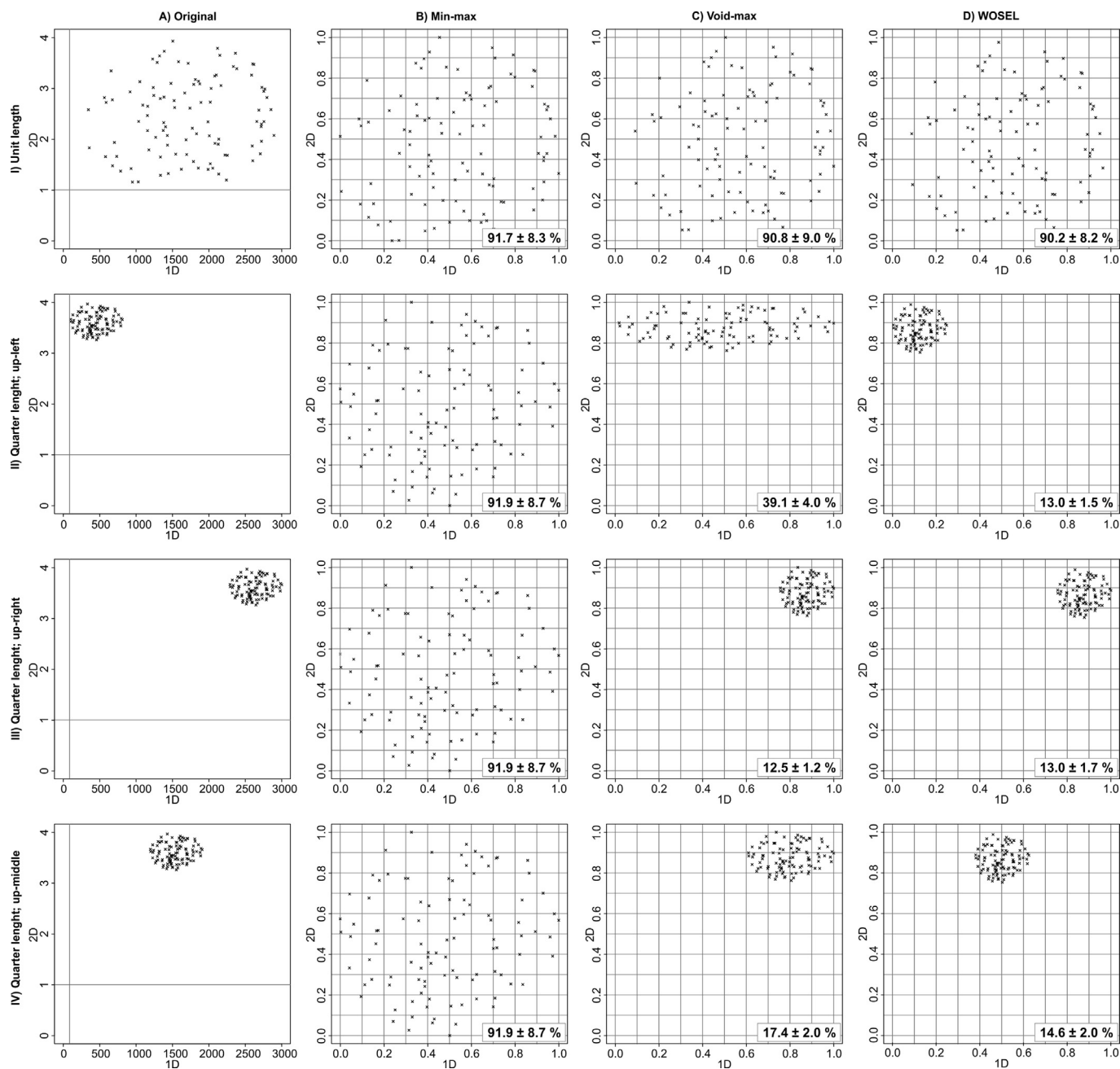
$$X_a = (Rt_1 - Rt_0)/(Rt_{\text{end}} - Rt_0) \quad (3)$$

Orthogonality calculations based on seven different approaches were performed: (1) Gilar's orthogonality concept (GA) [8], (2) the modeling approach (MA) [17], (3) the conditional entropy approach (CE) [11], (4) the asterisk equations approach (AE) [16], (5) the convex hull approach (CH) [14], (6) the nearest neighbor approach (NN) [13] and (7) the Pearson correlation approach (PC) [9]. All the algorithms were computed in the R software (v. 3.3.0, <https://www.r-project.org>). In CH, the area of the usable separation space was set accordingly to Min-Max, Void-Max and WOSEL to obtain results analogous to those obtained with the other approaches. In the cases where the discretized methods were used, the normalized data for every method was divided into bins in a standardized manner, and the sum of the bins was correlated to the size of a data set (number of peaks). The number of equidistant bins in each dimension was set to be equal to the nearest lower integer of the square root of the sum of the peaks, resulting in a different bin count and size for given chromatogram, e.g. 10 bins of the size 0.1  $\times$  0.1 in each dimension (in the case of models with 100 data points), 50 of 0.02  $\times$  0.02 (2500-data-point model) and 27 of 0.037  $\times$  0.037 (769 features; Fig. 3, example I). For Gilar's approach, the orthogonality was computed *via* Eq. (4), which was improved by Watson et al. [12] for multicomponent real samples, and where the number of peaks is not a square value:

$$O = \sum \text{bins} / (1 - \exp(-\alpha)) P^2, \quad (4)$$

and where  $P^2$  is the total number of bins spanning the normalized retention space,  $\sum \text{bins}$  is the count of bins occupied by a compound retention time and  $\alpha$  is the ratio of the compound number to  $P^2$ .

The results from the artificial chromatograms (Figs. 1, S1) were further investigated with six examples of real separations (Figs. 3, 4) by means of comprehensive two-dimensional gas chromatography coupled with time-of-flight mass spectrometry (GC  $\times$  GC/TOF-MS).



**Fig. 1.** Four different scenarios of an artificial chromatogram containing 100 data points (A) and the impact of three different scaling approaches (B–D) on the appearance of the normalized separation space. The values in the bottom right-hand corner show the orthogonality value of the model and its standard deviation ( $\pm 2$  SD) for 50 models generated *via* the same algorithm based on random sampling.

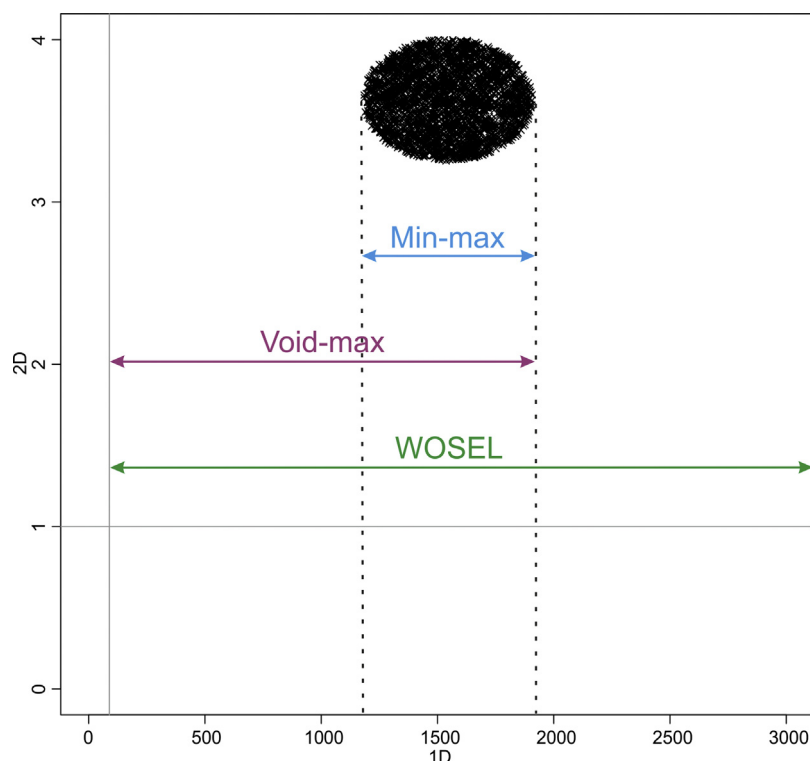
Ethoximated and trimethylsilylated human urine and plasma samples and one petroleum sample were used. The void times ( $Rt_0$ ) were calculated using LECO Column Calculator and LECO ChromaTOF-GC software (v. 4.51.6.0; LECO, St. Joseph, Michigan, USA). In the case of the artificial chromatograms, the range was set to 89 s and 3000 s for the first dimension and 1 and 4 s for the second one, mimicking a real GC  $\times$  GC analysis (Fig. 2). For real separations, true void time values based on specific column combinations and temperature programme were used.

### 3. Theory

The orthogonality of an analytical system towards a given sample type (e.g. human biofluid, environmental soil matrix) is

evaluated in the available separation space delineated by its beginning ( $Rt_0$  or  $Rt_{\min}$ ) and end ( $Rt_{\max}$  or  $Rt_{\text{end}}$ ). Therefore, the extent of the available separation space must be carefully defined [3]. This important parameter has already been addressed in the convex hull orthogonality approach [14,15], which does not use a scaling algorithm. The impact of the different extent of usable retention space on other approaches to orthogonality has not yet been determined.

The “normalization” of multisource data was introduced in order to compare the orthogonality of the information obtained with three different analytical techniques in drug analysis [5]. Direct comparison of the data was not possible, because the techniques differed in terms of their analysis time. Similar observations were made with the data from 2D chromatographic separations. In GC  $\times$  GC, for instance, the analysis time of the first dimension is



**Fig. 2.** Three types of scaling used for the calculation of the orthogonality. The void time in the given dimension (89 s and 1 s) is displayed as a grey line. Retention times are in seconds. The arrows depict the calculation for the first dimension. The second dimension calculation is analogical.

over three orders of magnitude longer than the secondary retention time.

Slonecker et al. [6] adopted this procedure to plot multisource data and compare its information content. However, they called it range-scaling transformation. This formulation is mathematically correct, since the transformation does not change the data distribution from non-Gaussian to normal, but changes the different magnitudes of the first and second dimensions into unit length. Therefore, we use the term “scaling” instead of “normalization” in this work.

Scaling for orthogonality calculations has been performed via two similar algorithms (Eqs. (1), (2)). Generally, the relative retention time of the analyte is divided by the total range of the analysis time. The analysis time is delimited by the beginning and the end of the analysis. The beginning of the analysis is set as a void time ( $Rt_0$ , Eq. (1)) – the retention time limit before no analyte can elute – or as the retention time of the first analyte to elute ( $Rt_{min}$ ) – the retention time limit before the peaks are out of interest (Eq. (2)); the end of the analysis as a retention time of the longest-eluting compound ( $Rt_{max}$ ).

When the retention time of peak apex of the last compound to elute is used, the evaluation of the orthogonality is sample- and separation-fluctuation- dependent. In other words, the same analytical system calculates different orthogonality values for each individual sample analyzed. This fact is caused by analytes with a high retention time in the first or second dimension occasionally being present in the sample. Therefore, Eq. (1) and Eq. (2) are effective for analyses targeted on selected group of compounds, where the retention and overall separation pattern is known. In “untargeted type” method, analyte could be anywhere before  $Rt_{min}$  or behind the  $Rt_{max}$  of particular sample used for the orthogonality-based system evaluation. WOSEL solves this problem employing the maximum available separation space delineated by the void time and the end of the method ( $Rt_{end}$ ) in both dimensions. The void time acts as the limit before no analyte can elute. However,

in particular separation methods the void time could be replaced by another parameter that unambiguously sets the start of the analysis, e.g. electroosmotic flow in capillary electrophoresis or solvent delay in GC. Theories of separation methods do not provide parameter defining end of the analysis or the separation space (analogical to  $Rt_0$ ) and traditionally  $Rt_{max}$  (retention coordinate of last peak) is employed. This value is easy to obtain, however does not reflect reality where the analysis ends. The data collection is usually stopped at the end of last peak of interest plus time required for retention fluctuations. The end of the method is thus set to enable the capture of all possible compounds of interest, regardless of their actual presence in the sample, and does not depend on a unique sample composition but generally on the sample type (e.g. urine, plasma, tissue, cells) and purpose of the analysis.

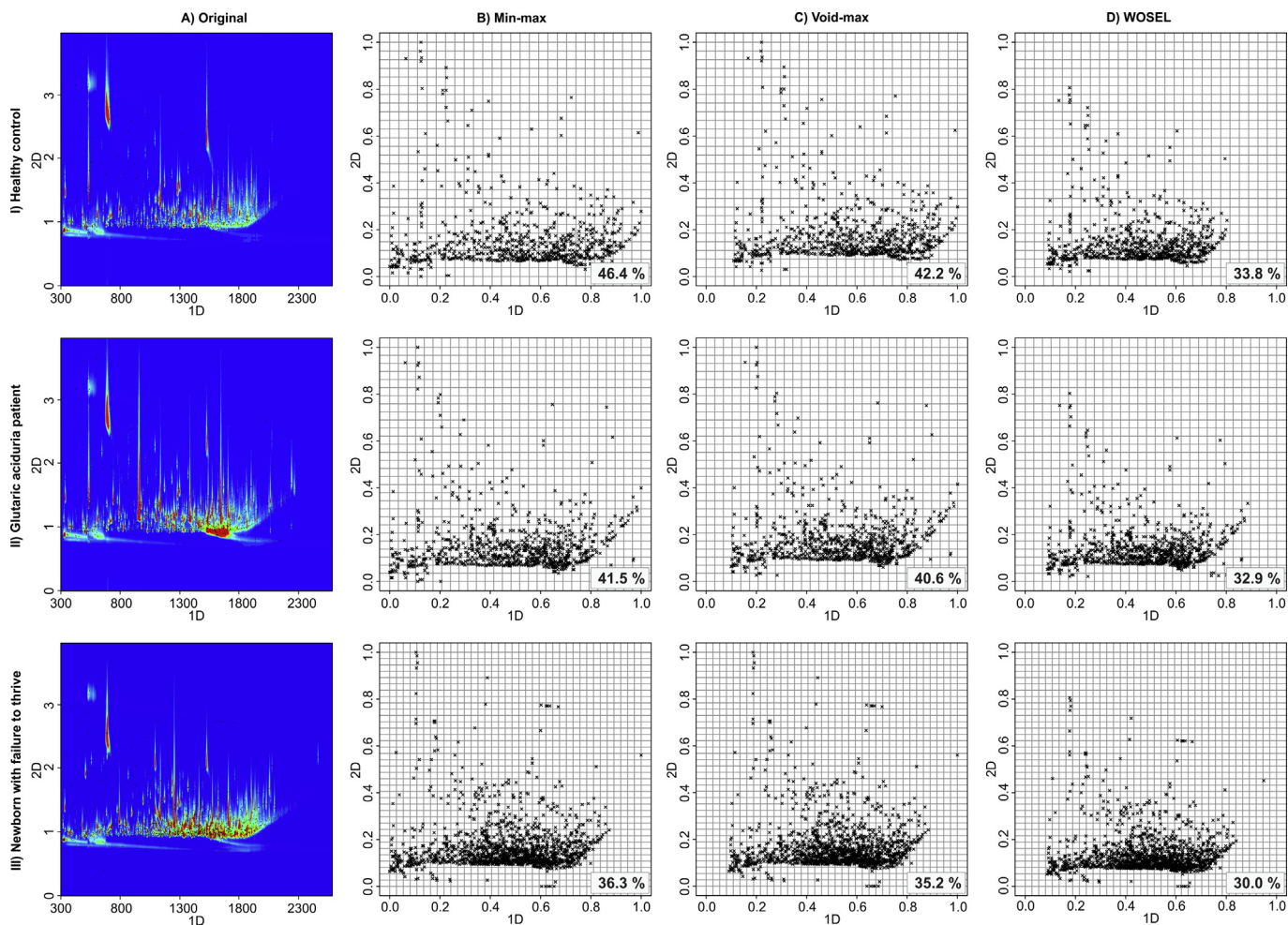
A specific form of method-based scaling was previously used by Rutan et al. [19] for LC  $\times$  LC separations, where normalization over the gradient time was proposed, as it represented the utilization of the space better. However, the separation of the analytes performed in the upper plateau space following the end of the gradient was not taken into account in this work.

Here we propose using the end of the method, since the proper analytical method is set to capture all the analytes potentially present in a given sample type. This option makes the orthogonality-based system optimization sample-independent and more robust.

#### 4. Results and discussion

The four parameters  $Rt_0$ ,  $Rt_{min}$ ,  $Rt_{max}$  and  $Rt_{end}$  distinguish the scaling approaches. For easy orientation among the approaches, we will call them Void-Max scaling (Eq. (1)), Min-Max scaling (Eq. (2)) and WOSEL (Eq. (3)). In general, the beginning of the separation space could be delineated by  $Rt_0$  or  $Rt_{min}$ . Those two parameters differ in terms of the ease with which they can be determined.  $Rt_{min}$  can be derived directly from every chromatogram, while, in order





**Fig. 3.** The impact of a highly-retained compound in the first dimension on the orthogonality value with different scaling approaches. (A) chromatograms of real samples: (I) urine of healthy child (769 features), (II) urine of pediatric patient suffering from glutaric aciduria (884 features), (III) urine of newborn with failure to thrive (1370 features), TIC:  $m/z$  35–650, colour palette: 0–5%; (B) Min-Max scaling; (C) Void-Max scaling; (D) WOSEL. The values in the bottom right-hand corner show the orthogonality.

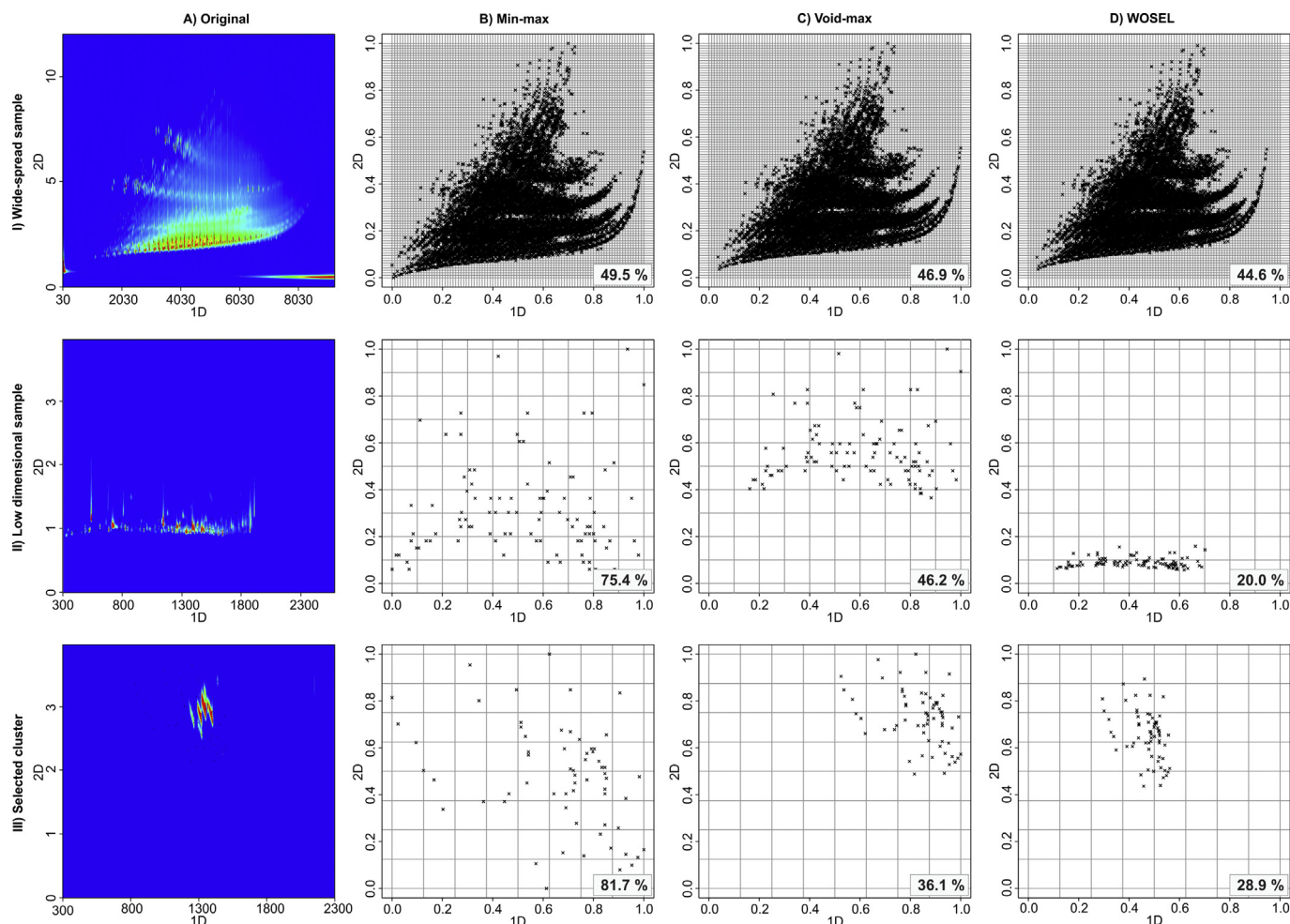
to evaluate  $Rt_0$ , additional analysis or computation based on the physicochemical properties of the system is required. For  $GC \times GC$ , Ryan et al. [3] suggested the use of alkanes for polar and alcohols for non-polar stationary phases. In LC, the choice of the unretained compound depends heavily on the separation and the detection system. As an example, an inorganic salt can be used as an unretained compound in reverse phase mode or ion suppression study can be used as an alternative approach. A void time calculator provides the simplest way of determining the void time, in spite of the fact that it is just a theoretical number and may differ slightly from the experimental value. However, the difference is, in terms of the determination of the available separation space, negligible.

The end of the separation space is easy to obtain from the chromatogram ( $Rt_{max}$ ) or from the parameters of the method ( $Rt_{end}$ ). In many application areas, such as toxicology, metabolomics, clinical diagnostics or environmental forensics, wide analytical windows must be set to enable the diagnostic analysis of highly retained analytes of interest that are just sporadically present in the sample. When an analytical method is developed, an appropriate analytical window is set in order to catch all the analytes of interest. As an example, in diagnosing mevalonic aciduria, two important diagnostic markers found are present: reduced levels of cholesterol and increased levels of mevalonic acid in urine. In  $GC \times GC$  analysis (regular column set BPX/BPX50), both of them are highly retained in the first (cholesterol, 1TMS) or second (mevalonolactone, 1TMS) dimension. Despite the fact that this metabolic disease is

very rare (approximately 30 patients have been reported), the analytical method for the diagnosis of metabolic disorders must be set to capture them [20].

In previous models artificial chromatograms of various shapes (e.g. rectangles, bands and diagonals) are used [10,16]. Data points in *in silico*-generated models used for the calculation of the impact of scaling on orthogonality values were designed to resemble the patterns observed in  $GC \times GC$  separations of human biofluids (Fig. S1). This profile of 2D peaks is composed of compounds with similar properties forming an oval cluster, as described by Giddings [1]. In 1D GC the compounds would form one broad peak; however in 2D GC elliptical profiles with fine structures (as a result of the increased separation power) are obtained. To mimic the reality, the impact of the scaling approach on the orthogonality value is also tested on models containing 2500 data points. This resembles the typical number of chromatographic peaks in  $GC \times GC$ -based metabolomics and is not prone to variability when a low number of data points is used, as reported by Watson et al. [12] and Camenzuli and Schoenmakers [16].

Gilar's orthogonality concept based on geometric surface coverage [8] was employed to demonstrate the impact of scaling on the orthogonality value (Figs. 1, 3, 4, S2, S3), since it is well known and widely discussed [4,10–12,16,17,19,21–23]. To assure that the conclusions are not inherent to GA, the three types of scaling (Void-Max, Min-Max and WOSEL) were also applied to other approaches to orthogonality: MA, CE, AE, CH and NN, and compared to PC,



**Fig. 4.** The impact of different types of scaling on real sample orthogonality evaluation. (A) chromatograms of real samples: (I) petroleum sample, 9185 features, TIC:  $m/z$  40–600, colour palette: 0–5%, (II) 100 organic acids in urine sample; EIC:  $m/z$  210–250, colour palette: 0–5%, (III) 64 sugar-like compounds in plasma, EIC:  $m/z$  217, 307, 319, 361, colour palette: 0–50%; (B) Min-Max scaling; (C) Void-Max scaling; (D) WOSEL. The values in the bottom right-hand corner show the orthogonality.

which does not require scaling. In CH, the area of the usable separation space was set accordingly to Min-Max, Void-Max and WOSEL to obtain analogous results and thus ones comparable to other approaches. These approaches were selected to represent each category of orthogonality computation.

Figs. 1, S2 and 4 show the impact of the exploitation of the separation space and different cluster locations on the  $x$ -axis on the orthogonality values obtained with three different scaling approaches. Figs. 1, S2 and S3 (scenario I) and Fig. 4 (example I) illustrate that the scaling approach has no significant impact on the orthogonality value when a sample widespread analytes is analyzed. On the other hand, if only a part of the separation space is covered by data points (Figs. 1 and S2, scenarios II–IV, Fig. 4, examples II and III) the orthogonality results obtained with the approaches differ profoundly. The same phenomenon was also observed when using four other orthogonality evaluation approaches (MA, CH, NN, AE; Tabs. S1, S2). Approaches comparing the informational similarity of both dimensions only behave differently. In the case of PC, it is not surprising, since the scaling algorithms are not implemented in the algorithm of the calculation. The results obtained with CE displayed a different trend. This interesting phenomenon caused by the application of informational theory (measurement of the random variable uncertainty) will be discussed in another paper (manuscript in preparation). The differences caused by different scaling approaches are more pronounced when the orthogonality is calculated from a higher number of input

data points (Fig. S2, Tab. S2). The cluster displacement in the second dimension (Fig. S3) invokes analogous results (Tab. S3). The standard deviations of the orthogonality results (Figs. 1, S2, S3; Tabs. S1–S6) document that *in silico* models give reproducible results. The models containing 100 data points (Figs. 1, S3; Tabs. S1, S3, S4, S6) showed greater variability of results than the 2500-point-models (Fig. S2, Tabs. S2, S5), as there were more possibilities for point translocation within the cluster area.

Min-Max scaling (Figs. 1 B, 3 B, 4 B, S2B, S3B) causes a severe misinterpretation of the real usage of the separation space usage. It enlarges the small quarter-length oval cluster into the whole normalized separation space, resulting in high orthogonality values (92% for 100 data points, 88% for 2500 data points). However, the result is not influenced by the localization of the cluster. The same effect is recognized on real GC  $\times$  GC separations (Fig. 4, examples II, III). Fig. 4B, example II shows that Min-Max scaling even masks poor GC  $\times$  GC method optimization (improper modulation period) or low sample dimensionality.

Void-Max scaling (Figs. 1 C, 3 C, 4 C, S2C, S3C) works well just for the cluster located at the end of the separation space (scenario III in Figs. 1, S2, S3). Both Void-Max and WOSEL scaling offer the same results (13% for 100 data points, 9% for 2500-data-point models). The altered location of the quarter-length cluster results in increased orthogonality with Min-Max but not with WOSEL scaling. Accordingly, Fig. 4, examples II and III show that Void-Max

scaling is the cause of the increase in the orthogonality value as compared to WOSEL.

Another important matter worth discussing is the relationship of orthogonality values and the analytical performance of different systems. Older scaling approaches overestimate orthogonality for chromatograms with low exploitation of the separation space. This phenomenon is obvious from Fig. 4, example III (Selected cluster), where Min-Max scaling shows a high orthogonality value of 81.7% because a very small portion (13.8%) of the real separation space is converted into a normalized one. This is caused by the fact that the extent of the real separation space that is converted into a normalized retention space is set according to the first and/or last compound of the sample to elute. Therefore, the space normalized by different samples represents diverse “portions” of the separation space and the sample-to-sample orthogonality results of the same analytical system are incomparable. The performance of the analytical system is thus optimized for one unique sample. The separation of another sample of the same type by the analytical system could be under-optimized.

Fig. 3 illustrates the benefit of WOSEL as compared to earlier scaling approaches – it does not require the evaluation of the approaches to orthogonality based on exactly the same sample or test mixture, because the end of the analysis is given by the end of the method instead of the retention time of the last compound to elute. The presence or absence of highly retained peaks thus has no impact on the extent of the real separation space (Fig. 3D). Finally, the orthogonality values are comparable for similar sample types, despite the fact that they differ in terms of the total number of features. Therefore, the analytical method optimized by the orthogonality value for the separation of metabolites in urine can also be used for the separation of plasmatic compounds. Fig. 3, examples I–III represents three urine samples that differ in their character and number of chromatographic peaks (769, 884, 1370). However, the orthogonality calculated by WOSEL varies in the narrow range of 30.0–33.8% compared to Min-Max scaling (36.3–46.4%) and Void-Max scaling (35.2–42.2%).

On the basis of this fact, we can conclude that when WOSEL is applied, the orthogonality value also reflects the perceived quality of the chromatogram. This conclusion is in accordance with the findings of Schure and Davis [4], who performed an extensive comparison of orthogonality metrics (20 orthogonality metrics and 55 combinations) and the perceived quality of chromatograms. They found that the correlation between the orthogonality metrics and grades of perceived quality depends on the data set. This phenomenon might be caused by the scaling approach. Moreover, they evaluated discretized metrics (such as GA) as offering a better reflection of the perceived quality when paired with CH results by multiplication into one final product. It is worth mentioning that CH does not employ the Min-Max scaling algorithm. Therefore, the multiplication of the GA result by the CH orthogonality value can act to correct the distortion of the normalized separation space caused by the use of Min-Max. However, the importance of global and local orthogonality metrics was omitted in this general suggestion. This issue will be addressed in subsequent work (in preparation).

## 5. Conclusions

WOSEL, a new scaling approach for the calculation of orthogonality, was introduced and compared to two earlier scaling approaches. On the basis of artificial chromatograms and real GC × GC/TOF-MS separations, it was shown that in the case of samples with high exploitation of the separation space WOSEL provides similar results to both previously used scaling approaches. However, in the case of partial usage of the separation space, the usual scaling methods calculate incorrect and significantly higher levels

of orthogonality. Min-Max scaling heavily distorts the real usage of the separation space, enlarges a small cluster into the whole normalized space, masks poorly optimized separations and classifies them as highly orthogonal. Void-Max scaling exaggerates the orthogonality of clusters distant from the end of the method. The proposed WOSEL approach provides results that reflect the real exploitation of the separation space the most. It enables the usage of any real sample for the estimation of the orthogonality, despite the fact that samples vary in composition and not all the components are known. Compared to Min-Max scaling, WOSEL requires the measurement or computation of the void time. However, at the expense of one additional analysis/calculation sample, independent scaling is gained. The reward is the determination of orthogonality without the potential heavy influence of the first or last sporadic peak in the particular analysis.

## Acknowledgements

The work was supported by GAČR15-34613L, NPU I (LO1304).

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at <http://dx.doi.org/10.1016/j.chroma.2017.06.076>.

## References

- [1] J.C. Giddings, Two-dimensional separations: concept and promise, *Anal. Chem.* 56 (1984) 1258A–1270A, <http://dx.doi.org/10.1021/ac00276a003>, 1264A passim.
- [2] C.J. Venkatramani, J. Xu, J.B. Phillips, Separation orthogonality in temperature-programmed comprehensive two-dimensional gas chromatography, *Anal. Chem.* 68 (1996) 1486–1492, <http://dx.doi.org/10.1021/ac951048b>.
- [3] D. Ryan, P. Morrison, P. Marriott, Orthogonality considerations in comprehensive two-dimensional gas chromatography, *J. Chromatogr. A* 1071 (2005) 47–53, <http://dx.doi.org/10.1016/j.chroma.2004.09.020>.
- [4] M.R. Schure, J.M. Davis, Orthogonal separations: comparison of orthogonality metrics by statistical analysis, *J. Chromatogr. A* 1414 (2015) 60–76, <http://dx.doi.org/10.1016/j.chroma.2015.08.029>.
- [5] W. Steuer, I. Grant, F. Erni, Comparison of high-performance liquid chromatography, supercritical fluid chromatography and capillary zone electrophoresis in drug analysis, *J. Chromatogr.* 507 (1990) 125–140, [http://dx.doi.org/10.1016/S0021-9673\(01\)84188-1](http://dx.doi.org/10.1016/S0021-9673(01)84188-1).
- [6] P.J. Slonecker, X. Li, T.H. Ridgway, J.G. Dorsey, Informational orthogonality of two-dimensional chromatographic separations, *Anal. Chem.* 68 (1996) 682–689, <http://dx.doi.org/10.1021/ac950852v>.
- [7] J.M. Davis, D.R. Stoll, P.W. Carr, Effect of first-dimension undersampling on effective peak capacity in comprehensive two-dimensional separations, *Anal. Chem.* 80 (2008) 461–473, <http://dx.doi.org/10.1021/ac071504j>.
- [8] M. Gilar, P. Olivova, A.E. Daly, J.C. Gebler, Orthogonality of separation in two-dimensional liquid chromatography, *Anal. Chem.* 77 (2005) 6426–6434, <http://dx.doi.org/10.1021/ac050923i>.
- [9] Z. Liu, D.G. Patterson, M.L. Lee, Geometric approach to factor analysis for the estimation of orthogonality and practical peak capacity in comprehensive two-dimensional separations, *Anal. Chem.* 67 (1995) 3840–3845, <http://dx.doi.org/10.1021/ac00117a004>.
- [10] M. Gilar, J. Fridrich, M.R. Schure, A. Jaworski, Comparison of orthogonality estimation methods for the two-dimensional separations of peptides, *Anal. Chem.* 84 (2012) 8722–8732, <http://dx.doi.org/10.1021/ac3020214>.
- [11] M.R. Pourhaghighi, M. Karzand, H.H. Girault, Orthogonality of two-dimensional separations based on conditional entropy, *Anal. Chem.* 83 (2011) 7676–7681, <http://dx.doi.org/10.1021/ac2017772>.
- [12] N.E. Watson, J.M. Davis, R.E. Synovec, Observations on orthogonality in comprehensive two-dimensional separations, *Anal. Chem.* 79 (2007) 7924–7927, <http://dx.doi.org/10.1021/ac0710578>.
- [13] W. Nowik, S. Héron, M. Bonose, M. Nowik, A. Tchaplá, Assessment of two-dimensional separative systems using nearest-neighbor distances approach. Part 1: orthogonality aspects, *Anal. Chem.* 85 (2013) 9449–9458, <http://dx.doi.org/10.1021/ac4012705>.
- [14] G. Semard, V. Peulon-Agasse, A. Bruchet, J.-P. Bouillon, P. Cardinaël, Convex hull: a new method to determine the separation space used and to optimize operating conditions for comprehensive two-dimensional gas chromatography, *J. Chromatogr. A* 1217 (2010) 5449–5454, <http://dx.doi.org/10.1016/j.chroma.2010.06.048>.

- [15] C. Cordero, P. Rubiolo, B. Sgorbini, M. Galli, C. Bicchi, Comprehensive two-dimensional gas chromatography in the analysis of volatile samples of natural origin: a multidisciplinary approach to evaluate the influence of second dimension column coated with mixed stationary phases on system orthogonality, *J. Chromatogr. A* 1132 (2006) 268–279, <http://dx.doi.org/10.1016/j.chroma.2006.07.067>.
- [16] M. Camenzuli, P.J. Schoenmakers, A new measure of orthogonality for multi-dimensional chromatography, *Anal. Chim. Acta* 838 (2014) 93–101, <http://dx.doi.org/10.1016/j.aca.2014.05.048>.
- [17] Z.-D. Zeng, H.M. Hugel, P.J. Marriott, A modeling approach for orthogonality of comprehensive two-dimensional separations, *Anal. Chem.* 85 (2013) 6356–6363, <http://dx.doi.org/10.1021/ac400736v>.
- [18] P.G. Stevenson, M. Mnatsakanyan, A.R. Francis, R.A. Shalliker, A discussion on the process of defining 2-D separation selectivity, *J. Sep. Sci.* 33 (2010) 1405–1413, <http://dx.doi.org/10.1002/jssc.200900779>.
- [19] S.C. Rutan, J.M. Davis, P.W. Carr, Fractional coverage metrics based on ecological home range for calculation of the effective peak capacity in comprehensive two-dimensional separations, *J. Chromatogr. A* 1255 (2012) 267–276, <http://dx.doi.org/10.1016/j.chroma.2011.12.061>.
- [20] D. Haas, G.F. Hoffmann, Mevalonate kinase deficiencies: from mevalonic aciduria to hyperimmunoglobulinemia D syndrome, *Orphanet J. Rare Dis.* 1 (2006) 13, <http://dx.doi.org/10.1186/1750-1172-1-13>.
- [21] B. Omais, M. Courtiade, N. Charon, J. Ponthus, D. Thiébaud, Considerations on orthogonality duality in comprehensive two-dimensional gas chromatography, *Anal. Chem.* 83 (2011) 7550–7554, <http://dx.doi.org/10.1021/ac201103e>.
- [22] D.N. Bassanese, B.J. Holland, X.A. Conlan, P.S. Francis, N.W. Barnett, P.G. Stevenson, Protocols for finding the most orthogonal dimensions for two-dimensional high performance liquid chromatography, *Talanta* 134 (2015) 402–408, <http://dx.doi.org/10.1016/j.talanta.2014.11.037>.
- [23] D. Li, C. Jakob, O. Schmitz, Practical considerations in comprehensive two-dimensional liquid chromatography systems (LCxLC) with reversed-phases in both dimensions, *Anal. Bioanal. Chem.* 407 (2015) 153–167, <http://dx.doi.org/10.1007/s00216-014-8179-8>.



Contents lists available at ScienceDirect

Analytica Chimica Acta

journal homepage: [www.elsevier.com/locate/aca](http://www.elsevier.com/locate/aca)

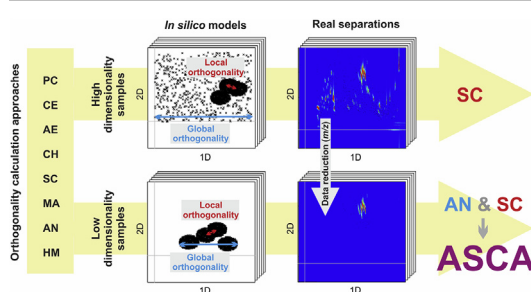
# Impact of sample dimensionality on orthogonality metrics in comprehensive two-dimensional separations

Jaroslava Jáčová<sup>a, b</sup>, Alžběta Gardlo<sup>a, b</sup>, Jean-Marie D. Dimandja<sup>c</sup>, Tomáš Adam<sup>a, b</sup>, David Friedecký<sup>a, b, \*</sup><sup>a</sup> Laboratory of Metabolomics, Faculty of Medicine and Dentistry, Palacký University Olomouc, Hněvotínská 5, 779 00, Olomouc, Czech Republic<sup>b</sup> Department of Clinical Chemistry, University Hospital Olomouc, I. P. Pavlova 6, 779 00, Olomouc, Czech Republic<sup>c</sup> School of Mechanical Engineering, Georgia Institute of Technology, 801 Ferst Dr. NW, 30313, Atlanta, United States

## HIGHLIGHTS

- The impact of sample dimensionality on orthogonality was evaluated.
- Models generated *in silico* containing 2500 peaks and real separation examples were used.
- The best local and global orthogonality descriptors were identified.
- ASCA, a combined local and global orthogonality descriptor, was defined and evaluated.

## GRAPHICAL ABSTRACT



## ARTICLE INFO

### Article history:

Received 8 December 2018  
 Received in revised form  
 4 March 2019  
 Accepted 8 March 2019  
 Available online 12 March 2019

### Keywords:

Required dimensionality  
 Global orthogonality  
 Local orthogonality  
 Combined descriptors  
 Arithmetic mean of nearest neighbor  
 Geometric surface coverage

## ABSTRACT

Orthogonality is a key parameter in the evaluation of the performance of a 2D chromatography-based separation system. Two different perspectives on orthogonality are determined: the extent of the separation space utilized (global orthogonality) and the uniformity of the coverage of the separation space (local orthogonality). This work aims to elucidate the impact of sample dimensionality (the number of separation processes involved) on orthogonality evaluation through the use of descriptors from seven different algorithms utilizing mutually different properties of a chromatogram: Pearson correlation, conditional entropy, asterisk equations, convex hull, arithmetic mean (AN) and harmonic mean of the nearest neighbor, and geometric surface coverage (SC). Artificial chromatograms generated *in silico* and real GC  $\times$  GC separations of diesel, plasma, and urine were used for the evaluation of orthogonality. The sample dimensionality has a deep effect on the orthogonality results of all approaches. The SC algorithm emerged as the best descriptor of local orthogonality samples of both low and high dimensionality, the AN algorithm on the global orthogonality of low-dimensionality samples. However, in the case of samples of high dimensionality, AN consistently indicated just the exploitation of the whole separation space; therefore, only local orthogonality is optimized by means of SC. Since no approach was able to monitor both global and local orthogonality as a single value, a new descriptor, ASCA, was developed. It

**Abbreviations:** (%O), orthogonality values in percentages; (2D), two-dimensional; (AE), asterisk equations approach; (AFID), alkali flame ionization detector; (AN), arithmetic mean of nearest neighbor approach; (ASCA), new combined orthogonality descriptor; (CE), conditional entropy approach; (ECD), electron capture detector; (EIC), extracted ion chromatogram; (GC  $\times$  GC), two-dimensional gas chromatography; (GC  $\times$  GC/MS), two-dimensional gas chromatography coupled to mass spectrometry; (HD), high-dimensional(ity); (HN), harmonic mean of nearest neighbor approach; (CH), convex hull; (LD), low-dimensional(ity); (*m/z*), mass to charge ratio; (MA), modeling approach; (MS), mass spectrometry; (NN), nearest neighbor approach; (O), orthogonality; (PC), Pearson correlation; (PCA), principal component analysis; (SC), geometric surface coverage; (TIC), total ion chromatogram; (WOSEL), scaling approach.

\* Corresponding author. Department of Clinical Chemistry, University Hospital Olomouc, I. P. Pavlova 6, 779 00, Olomouc, Czech Republic.

E-mail address: [david.friedecky@upol.cz](mailto:david.friedecky@upol.cz) (D. Friedecký).

<https://doi.org/10.1016/j.aca.2019.03.018>

0003-2670/© 2019 Elsevier B.V. All rights reserved.

combines the best global (AN) and local (SC) orthogonality algorithms by averaging, giving the same importance to data spread and crowding. ASCA thus provides the best estimation of orthogonality.

© 2019 Elsevier B.V. All rights reserved.

## 1. Introduction

Increasing requirements for the enhanced separation of complex samples have led to the evolution of two-dimensional (2D) chromatography-based techniques [1]. In order to optimize the effectiveness of these 2D separations, the concept of separation orthogonality was first introduced by Venkatramani et al. [2] Orthogonality, in the context of 2D chromatography, is the expression of the level of separation space that is effectively utilized for the resolution of the analytes in a sample. In a perfectly ordered 2D chromatogram, the entire separation space would be uniformly covered by peaks [3]. Moreover, compounds with similar chemical functionalities would be located in the same area of the bidimensional chromatogram and separated from compounds with other functional groups [4]. This fully orthogonal chromatogram is only attainable in theory because of constraints on the instrument's operation and the types of compounds (and matrix interferences) present in any given sample. Nevertheless, the metric of 2D chromatographic orthogonality is useful in comparing the effectiveness of 2D separations either in the developmental stage of a method or for quality control purposes.

A number of mathematical algorithms that can be classified into three different categories have been applied for the evaluation of 2D separation orthogonality. The first type of orthogonality methods characterizes the extent to which two separation mechanisms provide complementary information [2]. Mathematical descriptors (such as the coefficient of correlation, or conditional entropy), are used to evaluate the separation quality of the system [5,6]. The correlation coefficient method generates values that reflect the amount of information shared between the two separation dimensions, and these values are subsequently converted to orthogonality [7]. Highly correlated dimensions result in low separation efficiencies, and vice versa [2]. The conditional entropy method is based on information theory [8], and measures the uncertainty of a random variable as a function of its distribution in the separation space [6]. In this case, the result (orthogonality) provides information about the probability of a peak appearing at a particular retention time, considering the overall separation pattern.

The second type of orthogonality methods focuses on the coverage of the separation space. A convex hull (CH) encircles the area occupied by the peaks and compares it to the whole useable retention space [9]. The approach, based on geometric surface coverage [10], compartmentalizes the normalized retention space into rectangular bins and computes the ratio of occupied bins to all possible bins, divided by a constant (typically valued at 0.63) that corresponds to the probability of bins matching. The nearest neighbor approach (NN) [11] measures the distance from each peak in the 2D chromatogram to its nearest neighbor, and computes an average distance value for the overall separation. The asterisk equations concept (AE) [12] computes the spread of the peaks along each of four axes that cross the separation space and combines it into a final orthogonality value.

Finally, the third type of orthogonality methods, i.e., a geometric approach to factor analysis [5] and modeling approach [13], attempts to combine both peak dissemination in a separation space and information similarity by coupling the geometric surface

coverage to the peak spreading angle (inverse cosine of correlation) or the coefficient of determination ( $R^2$ ) into the final orthogonality value, respectively.

Schure and Davis [14] recently conducted a comparative study of orthogonality methods, and concluded that the best approaches should combine the “global” perspective of peak spreading with the “local” perspective of peak spacing. Chromatographic separations typically include areas in which peaks are well resolved and areas of higher congestion or clustering. Therefore, both local and global orthogonality must be optimized. However, there is no descriptor enabling the data crowding to be measured and the separation space to be exploited together. Schure and Davis [14] suggested combining the best global and local orthogonality approaches into one final product by multiplication.

The lack of uniformity in the peak distribution in chromatography is connected to the dimensionality of the sample, which was defined by Giddings [4] as the number of independent variables influencing the properties of the sample and thus its chromatographic behavior. One-dimensional samples consist of compounds of the same chemical class (e.g., saturated hydrocarbons). They do not exhibit any benefit when separated via a 2D system, since there is no other separation principle (than London dispersion forces) that can be involved. Compounds form a narrow band crossing the 2D separation plane. Two-dimensional samples contain compounds of two different functionalities (e.g., a mixture of saturated hydrocarbons and alcohols) that can be separated using two different mechanisms (e.g., London dispersion forces in the first dimension, hydrogen bonding in the second dimension). Such samples (called low-dimensionality samples, LD) have a dimensionality comparable to the dimensionality of the system. They provide ordered chromatograms in which the compounds form groups or lines of homologous series useful for the identification of unknowns [15].

However, Giddings [4] has predicted that the sample dimensionality of very complex samples is likely to be in the range of 2–3 orders of magnitude larger than the dimensionality of a 2D system. These samples of high dimensionality (e.g., urinary metabolome) contain compounds of many functional groups (e.g., organic acids, acyl glycines, acyl carnitines, polyols, amino acids, purines), which interact to a different extent with stationary phases of the first and second dimensions. Therefore, the final peak distribution in the 2D plane depends not only on the column combination [1] but also on the sample components, specifically on the strength of their interactions with the stationary phases. Moreover, the sample dimensionality could be reduced even further by the analyst through the selection of properties necessary for the purposes of the analysis. If there is an ideal selectivity of stationary phases [4], the sample components will be separated according to one property in the first dimension and another property in the second dimension; other functional groups will not interact with either the first or second dimension at all. In such a case, an ordered low-dimensional chromatogram is obtained. In real separations, other functional groups in the sample also contribute to the separation, but not equally. According to Giddings [4], contributions of weak interactions are negligible and the chromatograms look like low-dimensional (“pseudo-ordered” chromatograms). Intermediate interactions result in the formation of satellite peaks, while strong

interactions will cause a distortion of the structure of the chromatogram. As a result, the lack of discernible patterns in the peak profiles (“disordered separation pattern”) is observed in samples of high dimensionality (HD) [4]. Chromatograms look like scatterplots: the whole separation space is randomly covered by data points, regardless of the structural classes of the constituent compounds in the sample.

Previous work [14] has not been successful in finding a reliable tool to explain the variation in 2D chromatography data, possibly as a result of the fact that the sample dimensionality component was not taken into consideration. All the samples (peptides, corn seed extracts) examined there [14] might be considered as an LD. The authors evaluated CH as the best global orthogonality descriptor out of 20 that were tested, as best reflecting the perceived quality of a chromatogram. However, the orthogonality values of other approaches were distorted by the Min-Max scaling [16] that calculates the orthogonality of the sample in variable (and thus incomparable) retention spaces. In CH, the scaling was not used and therefore only CH correlated well with the perceived quality of the chromatograms. Moreover, the NN descriptors were disadvantaged by very low values (by five orders of magnitude) compared to the others [14].

The aim of this work is to find the best global and local orthogonality descriptors of high-dimensionality and low-dimensionality samples and to combine them into one final descriptor by means of an appropriate method. The descriptors are tested using *in silico* models, containing 2500 peaks, and investigated further using real chromatograms of diesel, human plasma, and urine. To overcome the above mentioned limitations, data treatment by WOSEL scaling [16] and specific normalization of the NN descriptors are applied in this work.

## 2. Materials and methods

### 2.1. Artificial chromatograms

Artificial chromatograms were created *in silico* using the R software (v. 3.3.0; R Foundation for Statistical Computing, Vienna, Austria; <https://www.r-project.org>). These simulated models represented a variety of ordered (well-separated clusters) or disordered chromatograms, as they are referred to by Giddings et al. [4] 2500 data points were generated into a given two-dimensional separation space ( $89\text{--}3000 \times 1\text{--}4 \text{ s}$ ) in four different types of scenarios: designed uniform distribution (each peak is placed exactly in the middle of the bin; Table 1), random uniform distribution (the separation space is randomly covered by data points;

Table 1), clusters (data points form clusters in the blank separation space; Table 2), and hidden clusters (data points form one cluster or more against the background of randomly distributed data points; Table 4). Each given scenario consisted of different model chromatograms (Tabs. 1, 2, and 4). Each random model was generated 50 times via the same algorithm based on random sampling to evaluate the validity of the model (standard deviation values). The minimal distance between data points in the normalized separation space was set at a value of 0.001 (this value corresponds to the ratio of adjusted retention times of the second and the first dimension in the original retention space). The details of the generation of the models and rationale for the selection of the models are described elsewhere [16].

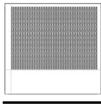


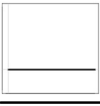
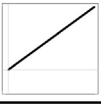
### 2.2. Real separation examples

Samples of human plasma or urine were obtained from routine analyses in our laboratory. Separations of diesel samples were provided by LECO Instrumente Plzeň, Czech Republic. In one set of runs, the same diesel sample was analyzed via four different methods by GC × GC/MS. In another set of runs, samples of human urine (Fig. 1) and plasma (Fig. 4) were analyzed on six different column sets by means of a GC × GC/MS system using a previously published method [17]. Details of the separation conditions and chromatograms are shown in the Supplementary Data. The retention coordinates of the compounds were taken from these runs for the orthogonality calculations.

### 2.3. Orthogonality calculations


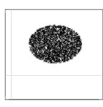
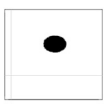
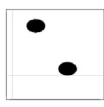
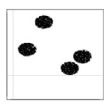
Orthogonality was evaluated by means of seven different approaches: Pearson correlation [5] (PC), conditional entropy [6] (CE), nearest neighbor [11] (NN), convex hull [9] (CH), asterisk equations [12] (AE), modeling approach [13] (MA), and geometric surface coverage [10] (SC). One representative approach from a group of orthogonality evaluation approaches with the same principle was chosen, since the approaches within a group are correlated [14]. The mathematical algorithms of the orthogonality approaches were converted into the R language (PC, CH, AE, MA, and SC) and MATLAB (CE, NN; MATLAB 2018a, Natick, Massachusetts, USA), in which all the results were calculated. A normalization step was performed via the WOSEL scaling approach [16]. The algorithms are provided in the Supplementary Data - Algorithms. In the case of the convex hull approach, the percentage of the separation space used was computed by a comparison of the area occupied by peaks via Delaunay's algorithm [9] to the useable separation space,

**Table 1**  
Orthogonality results [%] ( $\pm 2$  SD for random models) obtained with different orthogonality evaluation approaches for the determination of the orthogonality range of descriptors.

Scenario		A	B	C	D	E
Pattern						
Distribution		Designed	Random	Random	Designed	Designed
Approach	PC	100.0	98.8 $\pm$ 2.0	98.6 $\pm$ 2.2	96.5	0
	CE	100.0	85.8 $\pm$ 0.6	89.2 $\pm$ 0.8	–	0
	AN	156.3	99.9 $\pm$ 1.2	81.6 $\pm$ 1.1	3.1	4.4
	HN	156.3	67.3 $\pm$ 3.8	54.6 $\pm$ 3.6	3.1	4.4
	CH	96.0	99.2 $\pm$ 0.4	66.1 $\pm$ 0.3	0	0
	AE	95.9	96.3 $\pm$ 2.5	70.8 $\pm$ 0.8	0	0
	MA	158.7	100.3 $\pm$ 1.9	82.6 $\pm$ 2.0	3.1	0
	SC	158.7	100.3 $\pm$ 1.9	82.7 $\pm$ 2.0	3.1	3.2

**Table 2**

Cluster models in a blank separation space for evaluation of the orthogonality of LD samples. The average orthogonality value  $\pm 2$  SD in percentages is given for each scenario of 50 models by different orthogonality evaluation approaches. Scenario "E" consists of 54 models: 50 random + 4 designed clusters (51: corners, 52: diagonal, 53: line down, 54: line middle).

Scenario		A	B	C	D	E
Pattern						
Distribution		Random	Random	Random	Random	Random
Approach	AN	88.8 $\pm$ 0.5	54.1 $\pm$ 0.7	22.1 $\pm$ 0.1	25.7 $\pm$ 2.3	35.8 $\pm$ 4.3
	HN	59.7 $\pm$ 1.9	36.4 $\pm$ 2.6	14.8 $\pm$ 0.5	16.9 $\pm$ 1.6	23.0 $\pm$ 3.1
	CH	77.1 $\pm$ 0.1	27.8 $\pm$ 0.1	4.8 $\pm$ 0.0	11.4 $\pm$ 8.2	29.1 $\pm$ 26.8
	AE	78.3 $\pm$ 1.1	28.1 $\pm$ 0.7	4.9 $\pm$ 0.1	21.6 $\pm$ 24.1	43.6 $\pm$ 33.9
	SC	90.5 $\pm$ 0.9	44.8 $\pm$ 0.6	9.3 $\pm$ 0.1	11.7 $\pm$ 1.3	21.6 $\pm$ 3.7

delineated by the void time and the end of the analysis of the given dimension [16]. In the case of the SC approach, an equation (Eq. (1)) dedicated to the computation of the orthogonality of complex multicomponent samples was used [18]:

$$O = \sum bins / (1 - \exp(-\alpha)) P^2 \quad (1)$$

where  $P^2$  is the total number of bins spanning the normalized retention space,  $\sum bins$  is the count of bins occupied by a compound, and  $\alpha$  is the ratio of the compound number to  $P^2$ . The normalized separation space was divided by the equidistant bins. The number of equidistant bins in each dimension was set to be equal to the nearest lower integer of the square root of the sum of the peaks. It is worth mentioning that the denominator in Eq. (1) is for a smaller number of bins (about 100), usually expressed as  $0.63P^2$ . In this work a modified equation provided by Watson et al. [18] was employed.

Two approaches (PC and NN) did not provide orthogonality results in terms of percentages, and were thus not directly comparable to the others. To resolve this incompatibility, in the case of PC, the absolute correlation values were inverted (the absolute values were subtracted from the value of 1 [7]) and multiplied by 100%.

In the case of NN, the recalculation of harmonic and arithmetic means was developed (Eq. (2)):

$$O = 100\% \cdot \bar{x} \cdot (\sqrt{\sum peaks} - 1) / 0.64, \quad (2)$$

where  $\bar{x}$  is the harmonic or arithmetic mean,  $\sum peaks$  is the count of peaks resolved in the separation,  $(\sqrt{\sum peaks} - 1)$  is the number of distances in one dimension, and 0.64 is the fractional coverage achievable in a 2D space containing randomly dispersed compounds in terms of Poisson distribution [18]. Eq. (2) corrects both the main shortcomings of the NN approach [11]: (1) a different range of results and thus their incomparability to %O by other metrics [14]; (2) the dependency of the resulting values on the sum total of data points under consideration (e.g., an value of 0.0150 vs. 0.0029 for 100- and 2500-data point models, respectively [16]). Eq. (2) was developed for large data sets (2500 data points). It converts the absolute values of NN distances into percentages and thus makes the orthogonality results independent of the number of data points taken into account and comparable to other approaches. The value of 0.64 is derived from the Poisson law, which describes the probability of the occurrence of an event (e.g., a peak) in a given interval of space. However, for smaller datasets (up to about 100 data points) the value of 0.63 [18] should be used.

As the perceived quality of 2D chromatograms and the

calculated orthogonality values differ in many cases [14], the results were presented in both graphical (a picture of the dispersion of the retention coordinates in the separation space) and numerical (orthogonality value) formats. All the plots were constructed in Microsoft Excel 2016 (Microsoft, Santa Rosa, California, USA) and MedCalc (v. 16.4, MedCalc Software bvba, Ostend, Belgium; [https://www.medcalc.org](https://www.medcalc.org;); 2015). Multivariate data analysis techniques (principal component analysis (PCA), correlation analysis, matrix of distances, and cluster analysis) were applied, too. All the calculations were performed in the R software using home-made algorithms.

### 3. Results

When performing a two-dimensional separation technique, every analyst must face the problem of finding the best combination of separation principles to couple. The suitability of a given combination accounts for the challenges imposed by the dimensionality of the sample. There are several dozen orthogonality evaluation approaches based on different basic principles that answer diverse questions via the same descriptor (usually called orthogonality,  $O$ ). In order to find the best descriptor for a given sample dimensionality (LD or HD), artificial chromatograms of different scenarios (Tabs. 1, 2, and 4) were constructed. Considering the analytical power of current technologies, a higher number of data points (2500) was generated in these data sets when compared to previous works [3,10,12]. Moreover, the 2500-data point models also provide a substantial reduction in the variability of results based on random data sampling [18]. The conclusions of the *in silico* calculations were approved by the results obtained from real separations from different application areas.

The results of orthogonality evaluation depend heavily on the extent of the space, where the orthogonality is calculated. The space is defined by a scaling approach, as detailed in our previously published work [16]. Different scaling approaches offers diverse spaces for orthogonality evaluation. The WOSEL scaling approach was employed for all artificial and real chromatograms, since it is not dependent on the sample used for the optimization. In WOSEL, the separation space is delineated by the analytical method: it usually starts at the void time of the first and second dimensions and stops at the end of the data acquisition in the first dimension and the modulation period in the second dimension. However, the definition of the separation space can vary with the method. In some applications, the void time of the second dimension is also considered as a useable separation space. In these cases, the separation space in the second dimension ends at the void time of the second dimension, enabling the inclusion of peaks in the wrap-around.

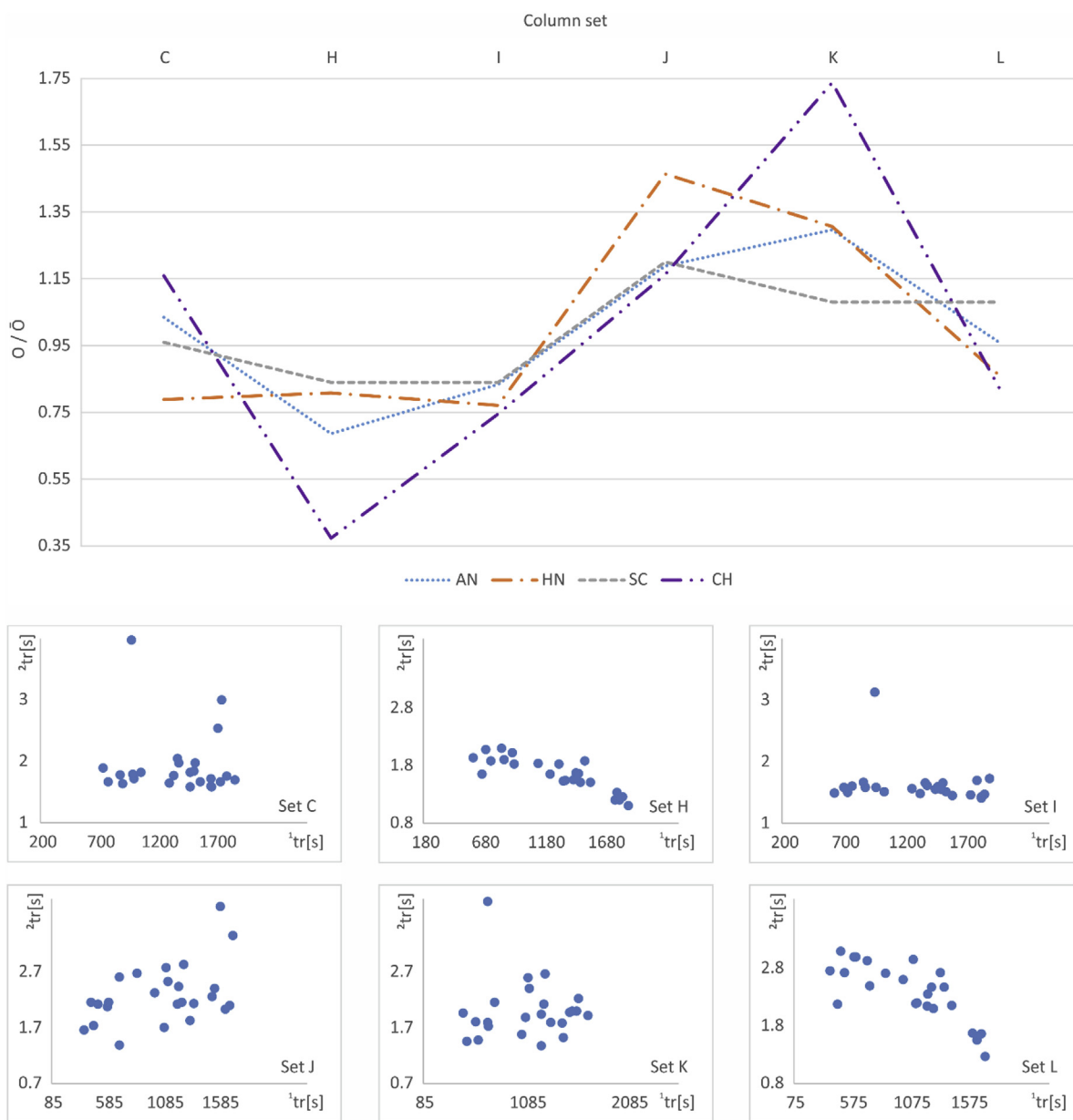


The orthogonality descriptors were investigated in three consecutive steps.

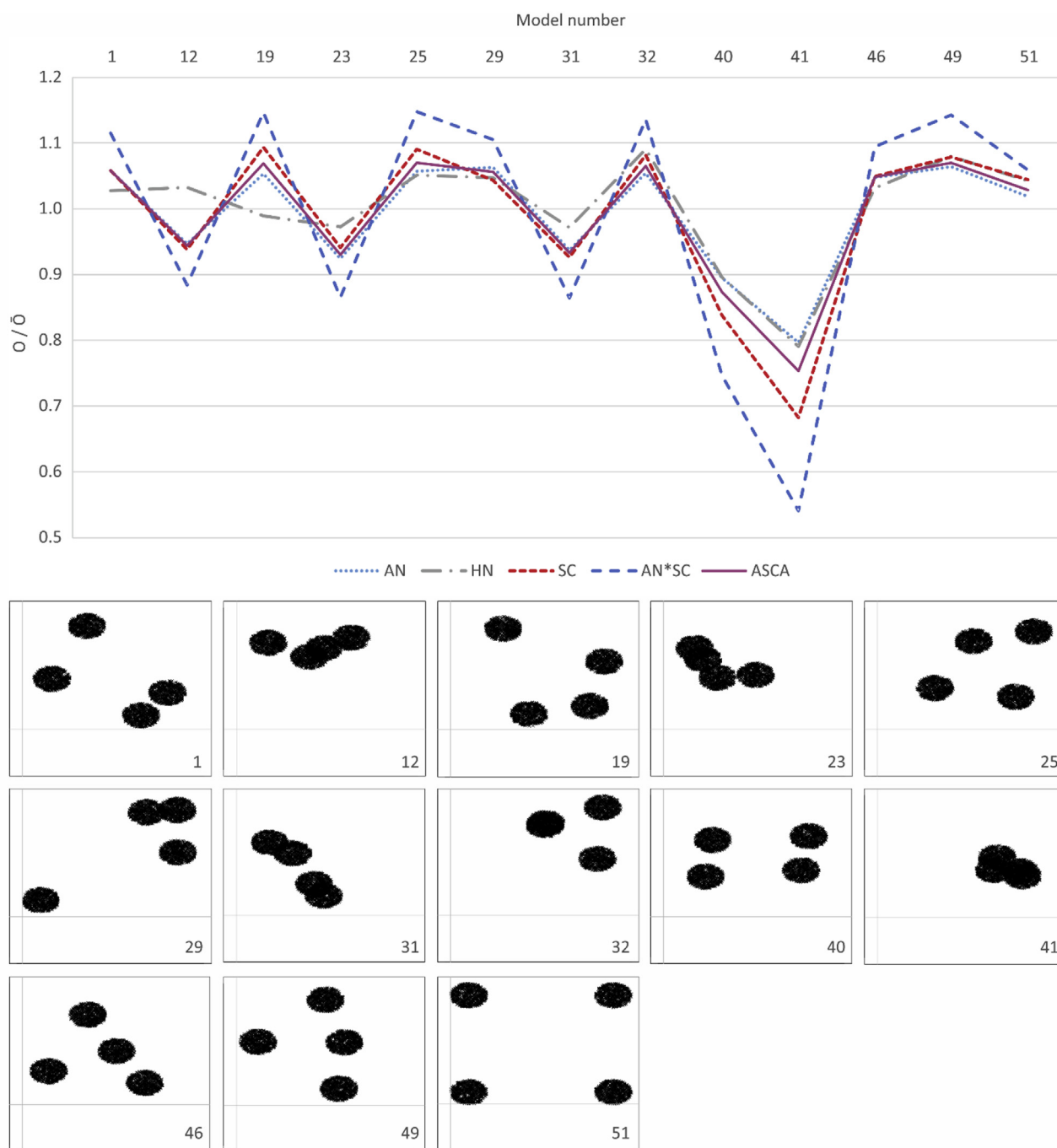
- (1) The ability to cover the entire orthogonality range from 0 to 100% was tested for all the approaches. The descriptors that failed were eliminated from further testing, because they were not able to distinguish properly between the best and the worst separations.
- (2) The whole range of orthogonality descriptors was applied to model data and real separation samples to find the most appropriate descriptor of local and global orthogonality for LD and HD samples. A combined global and local orthogonality descriptor was developed to overcome the challenge of making a selection based on either the maximal extent of the separation space used or the minimal clustering.

- (3) The sample dimensionality as an adjustable analytical parameter (based on the purpose of the analysis) was tested. The orthogonality of one sample was calculated with the approaches using the entire data set (HD) and a selected portion of the data set (LD), and the results were compared in terms of the perceived quality of the chromatograms.

In this work, the results are presented in both graphical (a picture of the dispersion of the retention coordinates in the separation space) and numerical (orthogonality value) formats (Tables 1–4 and Tables S1–S9, Fig. 4) and comparative charts (Figs. 1–3). To eliminate the differences between approaches in terms of the numerical values of the orthogonality (every approach has its own range of results), the %O obtained by the approaches were normalized to the average value for the given approach prior to the



**Fig. 1.** Dissemination of 25 urinary metabolites in the 2D separation space by GC  $\times$  GC/MS performed on six different column sets via the orthogonality descriptors: AN (arithmetic mean of nearest neighbor), HN (harmonic mean of nearest neighbor), CH (convex hull), SC (geometric surface coverage). Orthogonality value normalized to average value of the particular descriptor shown on the y axis.



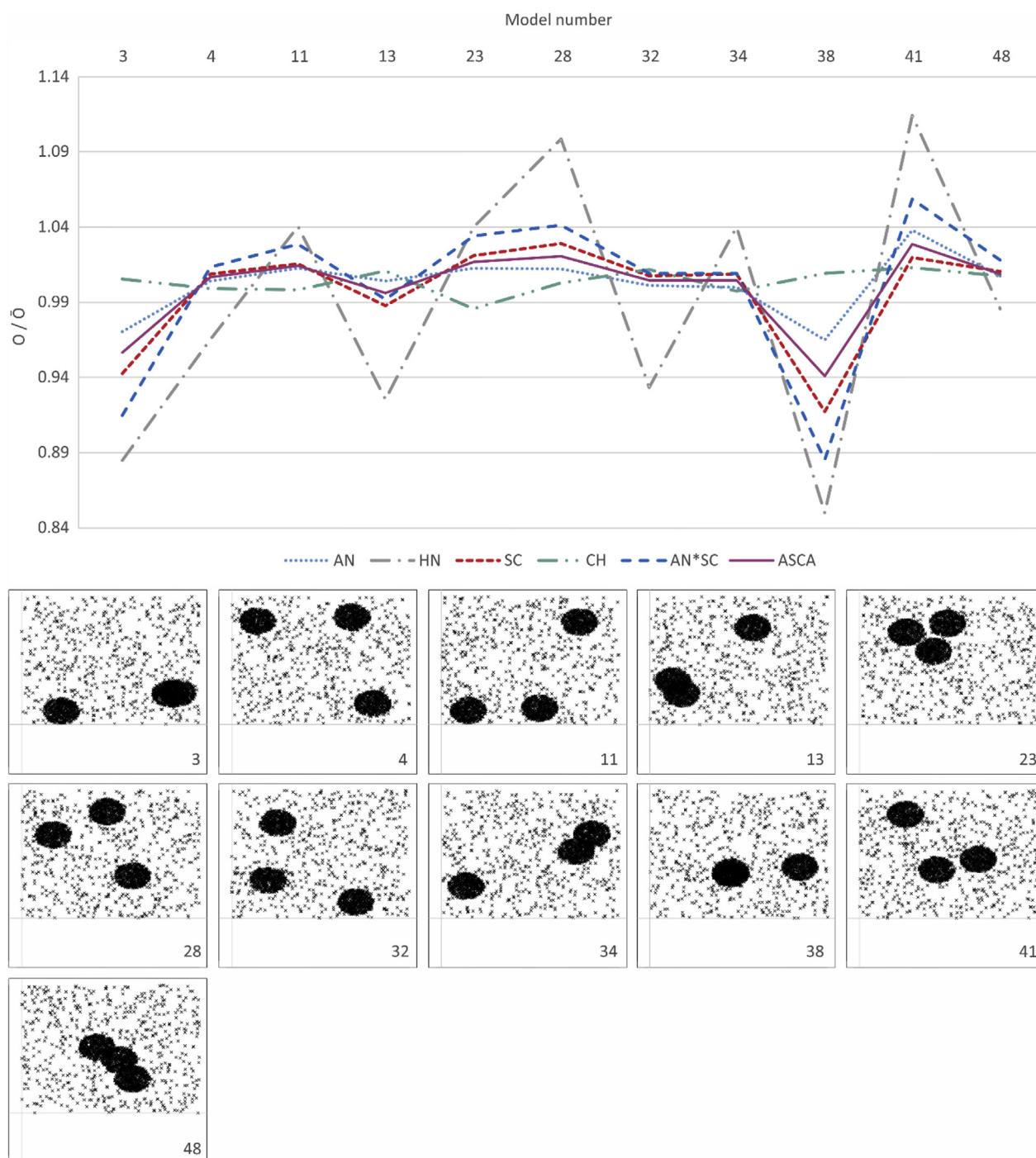
**Fig. 2.** Four-cluster models according to different orthogonality descriptors: AN (arithmetic mean of nearest neighbor), HN (harmonic mean of nearest neighbor), SC (geometric surface coverage), and a combination of the best two descriptors: AN\*SC (multiple of AN and SC), ASCA (geometric mean of AN and SC). Orthogonality value normalized to the average value of the particular approach shown on the y axis.

construction of the comparative plot.

### 3.1. Multivariate analysis of *in silico* models

In order to be consistent with the current practices in the comparison of orthogonality metrics [14], several multivariate data analysis techniques, including principal component analysis (PCA), correlation analysis, matrix of distances, and cluster analysis, were applied. However, these methods only took orthogonality values into account, regardless of the perceived quality of the chromatogram. All

these types of comparisons were therefore very challenging to make. For instance, the PCA of hidden three-cluster models showed differences between AE and CH and the group of AN, SC, and HN (Fig. S1). The artificial chromatograms were separated according to cluster overlay along the *x*-axes (based on the first principal component). However, the segregation is described by a combination of three different descriptors (AN, SC, and HN) that contribute to the first principal component predominantly and almost equally (Fig. S2). PCA enabled neither the determination of the best approach out of AN, SC, and HN nor the revealing of differences. Second



**Fig. 3.** Three-cluster hidden-cluster models according to different orthogonality descriptors: AN (arithmetic mean of nearest neighbor), HN (harmonic mean of nearest neighbor), SC (geometric surface coverage), CH (convex hull), and a combination of the two best descriptors: AN\*SC (multiple of AN and SC), ASCA (geometric mean of AN and SC). Orthogonality value normalized to the average value of the particular approach shown on the y axis.

principal component showed that CH evaluated a different property of the chromatograms. We were not able to determine which property of the chromatograms was reflected. However, we can conclude that CH is not able to detect cluster overlay.

### 3.2. Orthogonality range of descriptors in comparison

The orthogonality values calculated by different approaches

should cover the entire range (between 0 and 100%) and must distinguish between average and extreme cases, so we addressed this issue. Five different types of artificial chromatograms (shown in Table 1) were generated that spanned the extreme boundary conditions of orthogonalities between 0% and 100%, and also included one scenario with an intermediate value of % $O$  (the data were compressed in the second dimension by a factor of 1/3).

The high-end boundary condition ( $O = 100\%$ ) was investigated

through two different scenarios: designed distribution (Table 1, scenario “A”), where each data point is placed exactly in the middle of its bin, and random distribution (Table 1, scenario “B”). A comparison of the results of scenarios “A” and “B” indicates that all the approaches that were tested, excepting PC, AE, and CH, took the probability of peak occurrence at a given location into account, displaying lower values for random than designed distribution. All these approaches (except CE) reflect the Poisson law by dividing the calculated orthogonality value by 0.63. Therefore, the approaches (AN, HN, MA, and SC) displayed values above 100% for the designed distribution (column A). HN exhibits the highest decrease in orthogonality when the results of scenarios “A” and “B” are compared (Table 1). This phenomenon is caused by the fact that HN (developed to stress short distances) detects the proximity of randomly distributed data points. Therefore, it did not reach the value of 100% in the models with random distribution. With regard to CH, the probability constant is unimportant since this approach only reflects the extent of the separation space occupied by the peaks, not the peak distribution. Accordingly, there was a decrease in orthogonality with CH, observed when the peaks were placed at the center of the bin (scenario “A”) instead of random distribution (scenario “B”), which allowed the exploitation of the entire area of the bin. Surprisingly, AE did not reach an orthogonality of 100% either for the designed or for the random model. Moreover, analogously to CH, there was no decrease in orthogonality values for randomly distributed data points. However, compared to CH, AE is not dedicated to the quantification of the extent of the separation space used.

For the low-end boundary condition ( $O=0\%$ ), two scenarios, both mimicking 1D separations in the 2D plane, were generated: a horizontal line of peaks (Table 1, scenario “D”) and a diagonal line of peaks (Table 1, scenario “E”). Most of the approaches returned orthogonality values between 0 and 4%, which are acceptable results since the orthogonality does not behave ideally as a function of peak capacity [18]. However, the horizontal line model (scenario “D”) produced an orthogonality value nearing 100% for the PC; CE was not able to calculate the orthogonality of data points placed exactly in a line, since there is no uncertainty in  $Y$  in this case. In the case of simulated intermediate orthogonality (scenario “C”), both CE and PC offered the same orthogonality as for scenario “B”, despite the fact that the random data distribution was lowered in the second dimension by 1/3. The other approaches showed a decrease in orthogonality in scenarios “C” and “D” as compared to scenario “B”. Table 1 illustrates that CE and PC do not reflect the usage of the separation space. PC is, moreover, susceptible to the displacement of a single data point leading to a drastic change in the PC value, despite the overall pattern remaining the same (Fig. S3). It is worth mentioning that PC does not reflect the separation space coverage (Table 1, scenarios “C”, “D”, and “E”), so that all the approaches incorporating PC (e.g., correlation, coefficient of determination) into the calculation (e.g., MA) are essentially affected in the same way (Tabs. S1, S2). Therefore, PC, CE, and MA were not considered for further parts of this work.

### 3.3. Low-dimensionality samples

Low-dimensionality samples consist of analytes that can be separated into distinct groups of compounds with similar properties using two different displacement mechanisms [1]. The chromatograms of LD samples are characterized by the occurrence of group(s) of compounds in an empty separation space. Therefore, GC  $\times$  GC separations of diesel samples [19], polycyclic aromatic hydrocarbons [20], polychlorinated biphenyls [21], or peptides [14] are typical representatives of LD samples. To model the real separations of LD samples (Table 2), oval clusters of peaks in empty

separation spaces were generated to mimic the chromatographic behavior of compounds with a similar structure [16].

Table 2 exemplifies that all the approaches tested reflected the changes in the general trend (alternation of cluster diameter and count) of the scenarios. The descriptors offered different average orthogonality values caused by the quantification of the diverse property of coverage of the separation space [14]. AN characterizes orthogonality as the extent of the separation space uniformly covered by data points, HN as the number of data points gathered in clusters, SC as the fraction of bins occupied by peaks. AE shows high variability in the results per scenario. This phenomenon was most pronounced in scenario “E”, so the models of this scenario were explored in more detail (Tables S3, S4, Artificial chromatograms and the Original data source provided in the Supplementary Data). The asterisk equations approach combines a distribution of data points around four axes ( $Z+$ ,  $Z-$ ,  $Z1$ ,  $Z2$ ), crossing the separation space, into a final orthogonality value. Table S4 indicates that the orthogonality with AE does not correspond to the coverage of the separation space (cluster-to-cluster distance), but only reflects the distance of the clusters from the axes, especially from their intersection. So this approach overweights this feature (compared to cluster spacing), leading to the %O values of chromatograms with the accidental appearance of a cluster in the center of the separation space being heavily distorted (Table S3, model 51 vs. 38). This property of AE, its failure to achieve an orthogonality of 100%, and its not reflecting the random distribution (Table 1) lead us to exclude AE from the further evaluation of LD samples.

The most variable scenario, “E” (Table 2), was used to compare the ability of the orthogonality evaluation approaches to reveal differences between models. The orthogonality results obtained with different approaches are shown in Table S3, illustrating that CH provides an orthogonality value inflated close to the maximum (as high as 98.5%) for four clusters in the corners of a blank separation space. CH thus gives information about the span of the data and does not reflect the peak spacing [22]. As confirmed by the analysis of a urine sample by six different column sets, CH is strongly influenced even by the only peak of different chromatographic behavior, regardless of the uniformity in the properties of other compounds under the examination (Fig. 1, set “H” vs. set “I”).

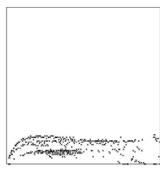
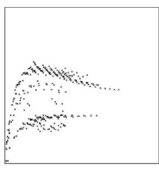
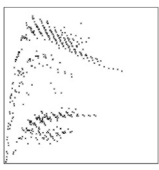
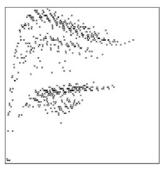
AN reflects the span of the exploited separation space the most, providing a higher orthogonality value for distant clusters (Fig. 2, Table S3, models 49 and 51 vs. 41 and 40). Compared to CH, AN represents the overall chromatographic pattern, not just the reflection of the only one different (highly-retained) peak (Fig. 1, set “H” vs. set “I”). However, when the average values of peak-to-peak distances are calculated, short distances between peaks in overlaid clusters are overweighted by the other long distances present in the chromatogram (Fig. 2, Table S3, model 1 vs. 29). Therefore, AN does not refer to data crowding or cluster overlap.

The mathematical algorithm used for the calculation of HN should amplify the impact of short peak-to-peak distances and suppresses the influence of long distances between clusters (Fig. 2, Table S3, model 23 vs. 49). However, this is not always the case, as can be seen in Fig. 2 and Table S3 (model 12 vs. 46).

SC calculates almost equal orthogonality values (about 23%, Table S3) for all the models of scenario “E” (four clusters with no overlap), regardless of the distances between them or location in the separation space. Values are reduced proportionally to the degree of overlap (Fig. 2, Table S3, models 25, 19, and 32 vs. 41, 40, and 31). Therefore, SC is able to detect the peak crowding (local orthogonality) but not the span of the separation space used (global orthogonality).

It can easily be seen that there is no descriptor that enables both crowding of peak apexes (local orthogonality metrics) and dispersion (global orthogonality metrics) to be monitored. Schure and

**Table 3**  
Evaluation of GC × GC/MS analyses of a diesel sample provided by four different analytical methods via four different orthogonalities. The charts depict the retention coordinates of the sample components. Details of the separation conditions are shown in the Supplementary Data.

Separation	A	B	C	D
Pattern				
AN	31.7	34.2	46.2	49.3
HN	21.1	16.7	23.3	19.6
SC	24.0	30.1	43.1	44.6
ASCA	27.8	32.1	44.6	46.9

Davis [14] suggested combining both the metrics by multiplication into one value. Here we have tested the combination of the most reliable global (AN) and local (SC) orthogonality metrics for LD samples. The multiplication of AN and SC only amplifies the trends in the behavior of the descriptors (Fig. 2). We propose here the averaging of SC and AN (“ASCA scoring”), postulating the equal importance of local and global orthogonality. Table S3 shows that AN provides higher %O values ranging from 28.5 to 41.3, SC from 17.7 to 23.6. Therefore, we recommend the usage of the geometric mean for SC and AN averaging. When the arithmetic mean is applied in this case, a higher value has a greater impact on the result and global aspects of orthogonality are emphasized. On the contrary, usage of the harmonic mean gives preference to local aspects of orthogonality by stressing the lower values.

The %O of four-cluster models according to ASCA ranges from 21.6 to 31.8%. All fully separated clusters were scored above 29.1%. ASCA correctly identifies the models perceived as the worst and the best chromatograms (model 41 and 51, respectively). For comparison, model 41, where all four clusters overlapped heavily, was correctly denoted as the worst case by all the approaches. On the contrary, the four clusters in the corners (model 51) were awarded the highest value only by ASCA and AN. In this case, SC proved misleading as a result of its amplifying the local aspect of orthogonality. However, compared to AN, ASCA revealed the cluster co-elution incorporating the information provided by SC (model 29). Therefore, ASCA reflects the influence of both the local and global orthogonality to the same extent.

Chromatograms of real LD samples show an apparent structure where compounds with similar functionalities form groups or lines (Table 3). It is obvious that both the extent of the exploited separation space and the peak spacing are indicators of the quality of the peak dispersion of the LD samples (Table 3). Analogously to LD models, AN and SC provide different information on the separation

in real samples. AN refers to a span of data points in the analysis providing the lowest value for the separation “A” and the highest value for the separation “D”. On the other hand, SC alone offers similar values for the separations “C” and “D”, indicating the clusters of data with similar structures (Table 3). However, when combined with AN in ASCA (with consideration also being given to the span of the separation space), analysis “D” is evaluated as being better than “C” (Table 3). The ASCA descriptor is thus beneficial.

### 3.4. High-dimensionality samples

High-dimensionality samples contain compounds of many functional groups that contribute in different degrees to the displacement provided by the first and second separation dimensions. Therefore, compounds do not form distinct groups; they are spread through the whole chromatogram [4]. HD samples do not present a frequent analytical problem as most of the complex samples (e.g., petrol) with a huge number of chemicals present are in fact of low dimensionality, since a low number of chemical types is present. Currently, there are only a few types of analyses which are able to deal with challenging HD samples (e.g., animal and human metabolomics [23], tobacco analysis [24]).

Models of HD samples were created as hidden-cluster models (Table 4) consisting of one or more clusters, each containing a quarter of the data points, placed among the randomly distributed data points. There is no ordered structure seen in the chromatograms and the data points are spread into the whole separation space. It is worth mentioning that there are many ways to construct HD models. In this work, a new cluster is made of points outside the clusters. However, we also created another type of model (a new cluster made up of points inside the clusters) that we evaluated in the same way with identical results (Tables S5–S7). The data are not discussed but are provided in the Supplementary Data.

**Table 4**  
Hidden-cluster models for evaluation of the orthogonality of HD samples. Average orthogonality value ± 2 SD in percentages for each scenario of 50 models by different orthogonality evaluation approaches. Each cluster contains a quarter of the data points.

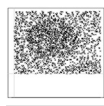
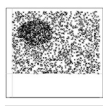
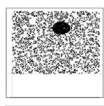

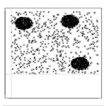
Scenario	A	B	C	D	E
Pattern					
Distribution	Random	Random	Random	Random	Random
Approach					
AN	98.1 ± 1.3	96.3 ± 0.9	92.8 ± 1.4	83.4 ± 1.7	71.1 ± 2.1
HN	63.1 ± 4.0	58.7 ± 3.1	46.1 ± 1.9	34.6 ± 3.5	27.7 ± 3.8
CH	99.0 ± 0.4	98.9 ± 0.5	98.9 ± 0.5	98.5 ± 0.7	97.3 ± 1.4
AE	87.4 ± 4.6	86.6 ± 7.9	85.9 ± 6.1	70.5 ± 13.4	59.8 ± 23.2
SC	96.0 ± 1.9	91.5 ± 1.8	85.5 ± 1.9	68.3 ± 1.6	47.8 ± 2.5

Table 4 illustrates that all the orthogonality evaluation approaches, with the exception of CH, can distinguish two general trends of the scenarios: a decrease in cluster diameter (scenarios “A”, “B”, and “C”) and the number of clusters (scenarios “C”, “D”, and “E”). However, CH consistently indicates only the exploitation of the whole separation space for all scenarios. As in the case of LD samples, AE also shows high variability and distortion of results (Tables 4, S8, S9). Therefore, CH and AE were excluded from other parts of this work.

The capability of the descriptors to distinguish tiny differences between particular models was studied in the most variable scenario, “E” (artificial chromatograms and the original data source provided in the Supplementary Data). The orthogonality results are shown in Table S8. HN showed the highest variability (and thus implies the highest sensitivity of the descriptor); nonetheless, this variability does not reflect the position and mutual separation of the clusters in a separation space. In Fig. 3 and Table S8 HN provided the same orthogonality values for a model containing distant, proximal, and co-eluted clusters (models 11, 23, and 34). On the contrary, a similar pattern (distant clusters) displayed different orthogonality values (model 28 vs. 32). This phenomenon was probably caused by the averaging of too many distances: every shortening of distance was overweighted by the extension of another one (for an explanation see Supplementary Data E1). Therefore, HN was excluded from further evaluation.

Fig. 3 and Table S8 reveals that AN does not reflect the cluster-to-cluster distance between hidden-cluster models. As an example, model 41 was evaluated as the best, although model 4 evidently depicts higher cluster-to-cluster distances. Moreover, co-eluted clusters showed the same orthogonality value as distant clusters (model 13 vs. 4). As mentioned above, that is because AN measures only the extent of the exploited separation space, which was almost the same for all the HD models (from 68.7 to 73.9%; Table S8). Considering local orthogonality, SC correctly showed the lowest value for two completely overlaid clusters (Fig. 3 and Table S8, model 38) and lower values for other overlaid cluster models (e.g., model 3). However, when the clusters were just in contact, there was no decrease in the %O observed (models 48 vs. 4), since the proportions of data points inside and outside the clusters remained unchanged.

The same phenomenon is present in real separations of certified plasma samples (Fig. 1, gray crosses), analyzed by GCxGC/MS on six different column sets. Neither AN nor HN was able to fully distinguish between the different separation patterns provided by the sets “A” and “B”. SC reflects the perceived quality of the chromatogram the best, providing the highest %O value for separation “F” and the lowest for “C”. There are certain aspects that one must be aware of when subjectively evaluating orthogonality on the basis of real HD data: just peak spacing can be optimized, since the entire separation space is filled by data points (it is not worthwhile to measure the extent of the separation space used). Moreover, in real HD samples, there is a substantial influence of a number of detected/displayed features on the perceived coverage of the separation space and if bleeding of the column is sometimes accidentally included in the peak list.

### 3.5. Required dimensionality

Biological samples consist of many compounds of different functionalities. Therefore, the structure of a two-dimensional chromatogram of such a sample is far from an ordered one. In untargeted metabolomics, all the compounds present in the sample are important; the sample behaves like an HD. On the contrary, in targeted metabolomics or clinical diagnostics, only a distinct group of compounds is examined; the sample is LD. Depending on the

purpose of the analysis, HD samples can also be transformed into LD ones. In the case of MS detection, an HD sample is convertible into an LD one by applying an appropriate  $m/z$  filter (generation of an extracted ion chromatogram, EIC). A decrease in sample dimensionality is also achievable by the usage of a specific detector (e.g., ECD, AFID).

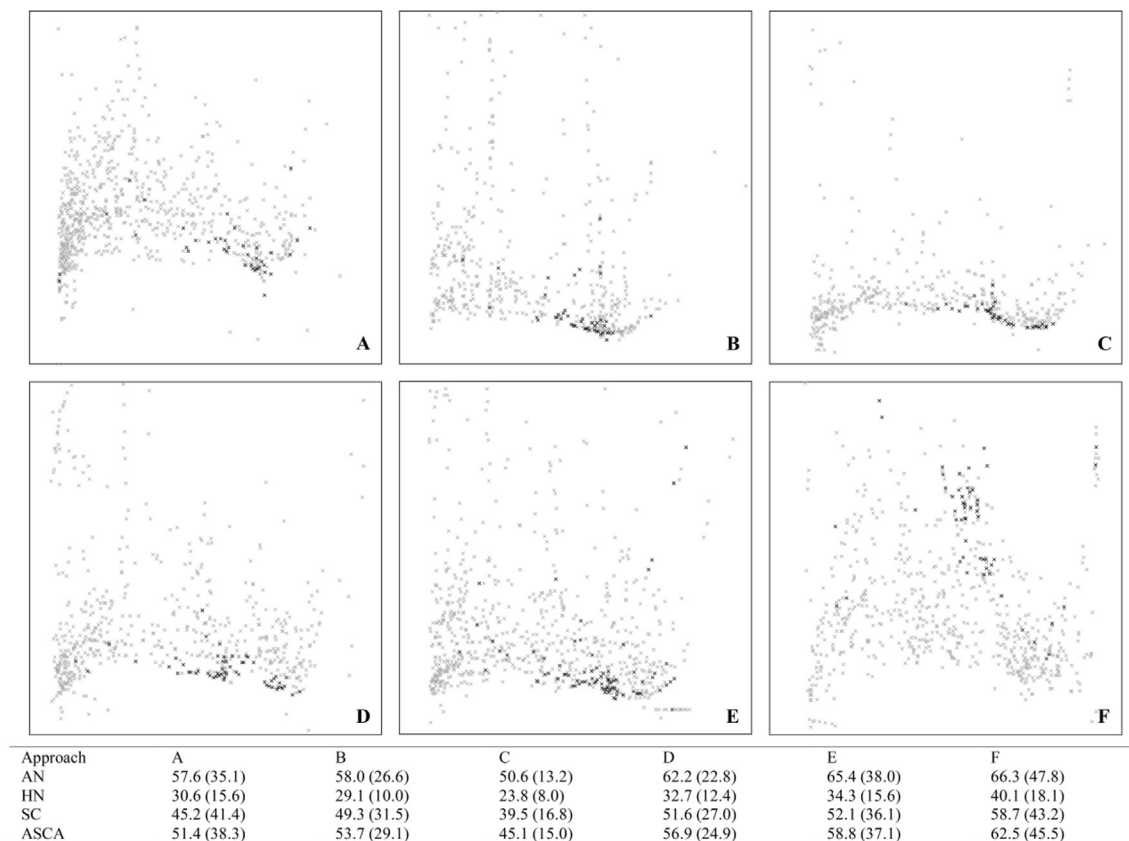
Fig. 4 shows GC × GC/MS separation of a certified plasma sample on six different column combinations (TICs in gray; EICs in black). TICs of the features show no ordering in the structure of the chromatogram, which is characteristic of HD samples. The application of an  $m/z$  217 filter (depicted in black) lowered the sample dimensionality and enabled the separation of sugar-based compounds according to their carbon number on the  $x$ -axis and functional groups (alditols, sugars, uronic acids) on the  $y$ -axis (Fig. S4). As compared to TICs of the features (Fig. 4, gray crosses), AN become informative after the dimensionality had been lowered (Fig. 4, black crosses), referring to the span of the separation space used. SC is useful in both cases, indicating data crowding. After the data reduction, ASCA meaningfully combines local and global orthogonality descriptors into one value, resulting in similar values for column sets “A” (lower fraction of data points in the cluster) and “E” (greater span of separation space).

## 4. Discussion

Orthogonality evaluation approaches provide different values in evaluation of the same chromatogram. These are caused by the different principles of the approaches, which give weight to different properties of the chromatogram. Orthogonality evaluation approaches usually fail to fit boundary conditions of %O = 0 and %O = 100. However, these are rarely seen in real separations. Therefore, the values provided by the orthogonality descriptors vary in different ranges and grades. A single %O value offered by the given approach is not informative. In the evaluation of a 2D separation system, comparison of several separations under different conditions must be performed. As a result, an improvement, stagnation, or decline of the property tested by the approach can be recognized. However, several orthogonality descriptors, such as PC or CE, reveal only the mathematical (in)dependence of the data and bear no relation to the perceived quality of the chromatogram. It is worth mentioning that some approaches, such as MA, that combine the PC value into the final %O should also be excluded from evaluation despite the fact that they provide acceptable results in most of our simulations.

Comparison of the %O results provided by different approaches is not easy to perform. Multivariate statistical methods are not very beneficial since they usually reduce the properties of the data set into main variables (known as principal components), regardless of the appearance of the chromatogram. It is again a purely mathematical operation that bears no relation to the chromatography. Therefore, in future comparisons of the performance of orthogonality descriptors, we recommend displaying chromatograms and the orthogonality values that are obtained together.

The proper choice of models to challenge the approaches is not an easy task. The appearance of real chromatograms depends on the separation mechanisms coupled together, the type of the sample, and the chemical properties of analytes of interest. In previous works (e.g. Refs. [3,12]), artificial chromatograms covered by rectangles, lines, bands, and stripes were used. Rectangles are unrealistic from the chromatographic point of view, while bands and lines are observed only in underoptimized separations. Therefore, ellipsis-shaped clusters rather than rectangles were generated. The same shape was initially used by Giddings [4]. Clusters like that are observed in separations of biological and environmental samples, but not in the analysis of petrol. Since the



**Fig. 4.** Comparison of GC  $\times$  GC/MS analyses of human plasma on six different column sets (A–F) via different orthogonality descriptors. TICs (untargeted analysis) of features are depicted in gray, EICs  $m/z$  217 (targeted analysis of sugar-based compounds) in black. The table below the picture shows %O provided by the descriptors for TIC and EIC. Details of the sample preparation and separation conditions are shown in the Supplementary Data.

models are always somewhat different from real samples, we have verified the results with separations of plasma, urine and diesel.

Moreover, there are many limitations that distort the orthogonality results and their interpretation. The scaling, as one of them, was discussed earlier [16]. Peak wrap-around is another important issue that needs to be considered when present. As mentioned above in “Results”, the dead time in the second dimension can be filled with separated sample components, especially when short modulation periods are applied. This can be treated by proper adjustment of the WOSEL scaling approach. However, more pronounced wrap-around can not only fill the “inaccessible” separation space of the void time of the second dimension but also increase the randomness of the peak distribution in the entire separation plane and “pull apart” peak clusters. This could be advantageous from the orthogonality point of view (especially when looking only at peak apices) and offer biased results. But this phenomenon should be prevented and eliminated by proper optimization of the analytical method and monitored using another analytical parameters (e.g. reduced peak capacity because of the greater widths of the wrap-around peaks).

The sample dimensionality also has to be considered. Most descriptors perform well for samples of low dimensionality, while separations of high dimensionality represent a more challenging task. Moreover, since no descriptor is able to refer to both global and local orthogonality, a combination of descriptors is beneficial. Our ASCA combined descriptor represents a more accurate alternative to the others and offers reliable results that are useful in the optimization of separation effectiveness, mainly for samples of low dimensionality or targeted data. In samples of high dimensionality

only local orthogonality via SC can be optimized, since the entire separation space is always used. However, the ASCA descriptor can also be applied, since the whole separation space is exploited: AN provides a constant value that (when averaged with SC) has no impact on the ranking of chromatograms.

Finally, it is worth mentioning that there is no sharp limit between samples of high and low dimensionality, with the exception of mixtures of pure standards. The dimensionality of real samples is usually very high (up to 10,000 [4]). However, the final distribution of the peaks in a 2D plane depends on the displacement mechanisms employed by the coupled dimensions or the adequacy of the separation system in general. As an example, if an improper column combination in GC  $\times$  GC is chosen, even 1D separation could be obtained. Sample dimensionality could thus be lowered by altering the analytical conditions, since the main demand of every analytical method is to achieve the goal of the analysis.

## 5. Conclusions

Orthogonality is a descriptor used in the evaluation of 2D separation systems. There are many algorithms for the calculation of orthogonality that provide different (and thus incomparable) orthogonality values for the same analytical system. Some of them (global orthogonality descriptors) refer to the extent of the separation space that is exploited, while others (local orthogonality descriptors) refer to the uniformity of apex-to-apex distances. In this work, seven different orthogonality calculation algorithms covering all the basic principles were tested. To set the importance of both separation dimensions as equal and make the

chromatograms comparable with each other, a WOSEL scaling approach was applied.

None of the descriptors that was tested was able to evaluate both perspectives of orthogonality. Geometric surface coverage was found to be the most reliable local orthogonality descriptor and arithmetic mean of nearest neighbor distances the most reliable global one. Combining them by averaging into a final orthogonality descriptor called ASCA made it possible to determine both orthogonality aspects by means of a single value.

Sample dimensionality, as an important property, was also considered in the evaluation of a multidimensional analytical system. Samples with the same dimensionality as the separation system (low-dimensionality samples) utilized a part of the separation space and both the global and local orthogonality aspects were important. In the case of high-dimensionality samples, the whole 2D chromatogram was covered by data points and the chromatograms differed only in the degree of clustering of the peak apexes. Samples with both dimensionalities (low and high) could be optimized by ASCA; however, in the high-dimensionality samples the global orthogonality component of the ASCA calculation is a constant value. It is worth to mention that not all of the sample properties are important for achieving the goal of the analysis.

Finally, it may now be time to change our thinking and we could consider the sample dimensionality as a parameter which is adjustable with regard to the goal of the analysis. Such a lowering of the sample dimensionality facilitates the optimization of the separation system and makes the results of the orthogonality calculation more explicit, important and useful, practical, and generally applicable.

## Declaration of interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

The work was supported by GAČR 18-12204S, NPU I (LO1304). We gratefully acknowledge Tomáš Kovalczuk from LECO Instrumente Plzeň, Czech Republic, for the use of the GC × GC analyses of diesel.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jaca.2019.03.018>.

## References

- [1] J.C. Giddings, Two-dimensional separations: concept and promise, *Anal. Chem.* 56 (1984) 1258A–1260A, <https://doi.org/10.1021/ac00276a003>, 1262A, 1264A passim.
- [2] C.J. Venkatramani, J. Xu, J.B. Phillips, Separation orthogonality in temperature-programmed comprehensive two-dimensional gas chromatography, *Anal. Chem.* 68 (1996) 1486–1492, <https://doi.org/10.1021/ac951048b>.
- [3] M. Gilar, J. Fridrich, M.R. Schure, A. Jaworski, Comparison of orthogonality estimation methods for the two-dimensional separations of peptides, *Anal. Chem.* 84 (2012) 8722–8732, <https://doi.org/10.1021/ac3020214>.
- [4] J.C. Giddings, Sample dimensionality: a predictor of order-disorder in component peak distribution in multidimensional separation, *J. Chromatogr. A* 703 (1995) 3–15, [https://doi.org/10.1016/0021-9673\(95\)00249-M](https://doi.org/10.1016/0021-9673(95)00249-M).
- [5] Z. Liu, D.G. Patterson, M.L. Lee, Geometric approach to factor Analysis for the estimation of orthogonality and practical peak capacity in comprehensive two-dimensional separations, *Anal. Chem.* 67 (1995) 3840–3845, <https://doi.org/10.1021/ac00117a004>.
- [6] M.R. Pourhaghighi, M. Karzand, H.H. Girault, Orthogonality of two-dimensional separations based on conditional entropy, *Anal. Chem.* 83 (2011) 7676–7681, <https://doi.org/10.1021/ac2017772>.
- [7] E. Van Gysegem, I. Crosiers, S. Gourvéneq, D. Massart, Y. Vander Heyden, Determining orthogonal and similar chromatographic systems from the injection of mixtures in liquid chromatography–diode array detection and the interpretation of correlation coefficients color maps, *J. Chromatogr. A* 1026 (2004) 117–128, <https://doi.org/10.1016/j.chroma.2003.10.116>.
- [8] P.J. Slonecker, X. Li, T.H. Ridgway, J.G. Dorsey, Informational orthogonality of two-dimensional chromatographic separations, *Anal. Chem.* 68 (1996) 682–689, <https://doi.org/10.1021/ac950852v>.
- [9] G. Semard, V. Peulon-Agasse, A. Bruchet, J.-P.P. Bouillon, P. Cardinaël, Convex hull: a new method to determine the separation space used and to optimize operating conditions for comprehensive two-dimensional gas chromatography, *J. Chromatogr. A* 1217 (2010) 5449–5454, <https://doi.org/10.1016/j.chroma.2010.06.048>.
- [10] M. Gilar, P. Olivova, A.E. Daly, J.C. Gebler, Orthogonality of separation in two-dimensional liquid chromatography, *Anal. Chem.* 77 (2005) 6426–6434, <https://doi.org/10.1021/ac050923i>.
- [11] W. Nowik, S. Héron, M. Bonose, M. Nowik, A. Tchaplá, Assessment of two-dimensional separative systems using nearest-neighbor distances approach. Part 1: orthogonality aspects, *Anal. Chem.* 85 (2013) 9449–9458, <https://doi.org/10.1021/ac4012705>.
- [12] M. Camenzuli, P.J. Schoenmakers, A new measure of orthogonality for multi-dimensional chromatography, *Anal. Chim. Acta* 838 (2014) 93–101, <https://doi.org/10.1016/j.aca.2014.05.048>.
- [13] Z.-D. Zeng, H.M. Hugel, P.J. Marriott, A modeling approach for orthogonality of comprehensive two-dimensional separations, *Anal. Chem.* 85 (2013) 6356–6363, <https://doi.org/10.1021/ac400736v>.
- [14] M.R. Schure, J.M. Davis, Orthogonal separations: comparison of orthogonality metrics by statistical analysis, *J. Chromatogr. A* 1414 (2015) 60–76, <https://doi.org/10.1016/j.chroma.2015.08.029>.
- [15] S. Zhu, X. Lu, K. Ji, K. Guo, Y. Li, C. Wu, G. Xu, Characterization of flavor compounds in Chinese liquor Moutai by comprehensive two-dimensional gas chromatography/time-of-flight mass spectrometry, *Anal. Chim. Acta* 597 (2007) 340–348, <https://doi.org/10.1016/j.jaca.2007.07.007>.
- [16] J. Jácová, A. Gardlo, D. Friedecký, T. Adam, J.-M.D. Dimandja, Sample-independent approach to normalize two-dimensional data for orthogonality evaluation using whole separation space scaling, *J. Chromatogr. A* 1511 (2017) 1–8, <https://doi.org/10.1016/j.chroma.2017.06.076>.
- [17] P. Wojtowicz, J. Zrostlíková, T. Kovalczuk, J. Schürek, T. Adam, Evaluation of comprehensive two-dimensional gas chromatography coupled to time-of-flight mass spectrometry for the diagnosis of inherited metabolic disorders using an automated data processing strategy, *J. Chromatogr. A* 1217 (2010) 8054–8061, <https://doi.org/10.1016/j.chroma.2010.09.067>.
- [18] N.E. Watson, J.M. Davis, R.E. Synovec, Observations on “orthogonality” in comprehensive two-dimensional separations, *Anal. Chem.* 79 (2007) 7924–7927, <https://doi.org/10.1021/ac0710578>.
- [19] J. V. Seeley, S.K. Seeley, E.K. Libby, J.D. McCurry, Analysis of biodiesel/petroleum diesel blends with comprehensive two-dimensional gas chromatography, *J. Chromatogr. Sci.* 45 (2007) 650–656.
- [20] C. Manzano, E. Hoh, S.L.M. Simonich, Improved separation of complex polycyclic aromatic hydrocarbon mixtures using novel column combinations in GC × GC/ToF-MS, *Environ. Sci. Technol.* 46 (2012) 7677–7684, <https://doi.org/10.1021/es301790h>.
- [21] J. Focant, A. Sjödin, D. Patterson Jr., Improved separation of the 209 PCBs using GCxGC-TOFMS, *Organohalogen Compd.* 66 (2004) 812–820. [http://www.dioxin20xx.org/ohc\\_database\\_search.htm](http://www.dioxin20xx.org/ohc_database_search.htm).
- [22] J. Leonhardt, T. Teutenberg, G. Buschmann, O. Gassner, T.C. Schmidt, A new method for the determination of peak distribution across a two-dimensional separation space for the identification of optimal column combinations, *Anal. Bioanal. Chem.* 408 (2016), <https://doi.org/10.1007/s00216-016-9911-3>.
- [23] N.P. Vasquez, M. Crosnier de bellaistre-Bonose, N. Lévêque, E. Thioulouse, D. Doummar, T. Billelte de Villemeur, D. Rodriguez, R. Couderc, S. Robin, C. Courderot-Masuyer, F. Moussa, Advances in the metabolic profiling of acidic compounds in children’s urines achieved by comprehensive two-dimensional gas chromatography, *J. Chromatogr., B* 1002 (2015) 130–138, <https://doi.org/10.1016/j.jchromb.2015.08.006>.
- [24] J. Vial, B. Pezous, D. Thiébaud, P. Sassiati, B. Teillet, X. Cahours, I. Rivals, The discriminant pixel approach: a new tool for the rational interpretation of GCxGC-MS chromatograms, *Talanta* 83 (2011) 1295–1301, <https://doi.org/10.1016/j.talanta.2010.07.059>.



# ŽIVOTOPIS

## OSOBNÍ ÚDAJE

---

Jméno: Jaroslava Jáčová  
Adresa: Praskova 446/5, 779 00 Olomouc  
Email: jaroslava.jacova@gmail.com

## VZDĚLÁNÍ

---

2013 – současnost	Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Katedra analytické chemie, doktorské studium oboru Analytická chemie
2016	Institut postgraduálního vzdělávání ve zdravotnictví – IPVZ, Akreditovaný kvalifikační kurz Odborné zdravotnické laboratorní metody
2012	Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Katedra analytické chemie, státní rigorózní zkouška Rigorózní práce: „Stanovení lanthanu v křemičitém skle“
2004 – 2006	Ostravská univerzita, Pedagogická fakulta, Centrum dalšího vzdělávání, Doplnující pedagogické studium pro učitele odborných předmětů. Závěrečná práce: „Hodnotová orientace žáků 2. ročníku SUPŠ sklářské“
1999 – 2001	Univerzita Pardubice, Fakulta chemicko-technologická, Katedra biologických a biochemických věd, magisterské studium oboru Analýza biologických materiálů Diplomová práce: „Adaptace lidských biologických hodin k přechodu na letní čas“
1996 – 1999	Univerzita Pardubice, Fakulta chemicko-technologická, Katedra biologických a biochemických věd, bakalářské studium oboru Klinická biologie a chemie Bakalářská práce: „Melatonin – funkce v lidském organismu, metody stanovení“

## STÁŽE

---

1. 6. – 31. 8. 2017 Neuroimunologický ústav, Slovenská akadémia vied, Bratislava, Slovenská republika

## PROJEKTY

---

2019	Molekulární podstata a diagnostické markery vybraných hematologických a systémových chorob X., UPOL, IGA_LF_2019_006
2018 – nyní	Charakterizace lidského lipidomu a metabolomu pro personalizovanou zdravotní péči a hledání biomarkerů: studie rakoviny ledvin, GAČR, projekt č. GA18-12204S
2018	Molekulární podstata a diagnostické markery vybraných hematologických a systémových chorob IX., UPOL, IGA_LF_2018_010
2015 – 2017	Statistika v metabolomice pro výzkum biomarkerů v medicíně, GAČR, projekt č. GF15-34613L
2014 – 2019	Podpora udržitelnosti Ústavu molekulární a translační medicíny, MŠMT, projekt č. NPU I LO1304
2012 – 2013	Soubor materiálů k badatelským aktivitám žáků ZŠ a SŠ v přírodních vědách, MŠMT, projekt č. CZ.1.07/1.1.00/26.0032
2008 – 2010	Průmyslový výzkum agregace skelných materiálů a jejich povrchové zpevnění pro využití v interiéru, MPO ČR, projekt č. FI-IM5/050
2006 – 2009	Impuls MSK, MPO ČR, projekt č. FI-IM3/177
1999 - 2001	GLASS CENTRUM Valašské Meziříčí, MPO ČR, projekt č. FB-C2/96

## PEDAGOGICKÁ ČINNOST:

---

2018 – nyní	vedoucí diplomové práce: „Analýza mastných kyselin s krátkým řetězcem v plasmě pacientek s gestačním diabetem“, Bc. Eliška Ivanovová, Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Laboratoř růstových regulátorů
2018	oponentka bakalářské práce: „Cílená a necílená metabolomická analýza krevní plazmy transgenních potkanů s tauopatií“, Lenka Kocurová, Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Laboratoř růstových regulátorů
2017 – 2018	vedoucí diplomové práce: „Imobilizovaná ureasa pro metabolomiku lidské moči založenou na GC/MS“, Bc. Nikola Čipáková, Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Katedra biochemie
2017 – 2018	vedoucí bakalářské práce: „Vliv pH na aktivitu ureasy pro GC/MS metabolomiku lidské moči“, Eliška Ivanovová, Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Laboratoř růstových regulátorů
2017	oponentka diplomové práce: „Proteasy a jejich využití v medicíně Optimalizace imobilizace ureasy“, Jitka Hutařová, Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Katedra biochemie

2015 – 2016	vedoucí bakalářské práce: „Analýza sacharidů a polyolů v moči pomocí GC/MS pro diagnostiku metabolických poruch“, Nikola Zdráhalová, Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Katedra fyzikální chemie
2013 – 2014	vedoucí diplomové práce: „Analýza metabolomu lidské moči pomocí GCxGC-TOF“, Bc. Marie Zdráhalové, Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Katedra biochemie
2003 – 2013	učitelka odborných předmětů, Střední umělecko-průmyslová škola sklářská Valašské Meziříčí

## **PUBLIKACE V ČASOPISECH S IMPAKTNÍM FAKTOREM**

---

2019	J. Jáčová, M. Jořenek, K. Pospíšková, L. Najdekr, L. Zajoncová, D. Friedecký, T. Adam, Urease-immobilized magnetic microparticles in urine sample preparation for metabolomic analysis by gas chromatography-mass spectrometry, <i>J. Chromatogr. A.</i> , v recenzním řízení
2019	J. Jáčová, A. Gardlo, J.-M.D. Dimandja, T. Adam, D. Friedecký, Impact of sample dimensionality on orthogonality metrics in comprehensive two-dimensional separations, <i>Anal. Chim. Acta.</i> (2019). doi:10.1016/j.aca.2019.03.018.
2017	J. Jáčová, A. Gardlo, D. Friedecký, T. Adam, J.-M.D. Dimandja, Sample-independent approach to normalize two-dimensional data for orthogonality evaluation using whole separation space scaling, <i>J. Chromatogr. A.</i> 1511 (2017) 1–8. doi:10.1016/j.chroma.2017.06.076.

## **OSTATNÍ PUBLIKACE**

---

2015	L. Bartáková, P. Daniš, J. Jáčová, Lukáš Müller, <i>Náměty k mimoškolní činnosti - chemie</i> , Univerzita Palackého v Olomouci 2016, ISBN 978-80-244-4745-2
2015	L. Bartáková, P. Daniš, J. Jáčová, L. Müller: <i>Laboratorní a terénní cvičení - chemie</i> , Univerzita Palackého v Olomouci 2015, ISBN 978-80-244-4590-8
2015	L. Bartáková, P. Daniš, J. Jáčová, L. Müller: <i>Čítanka k přírodním vědám - chemie</i> , Univerzita Palackého v Olomouci 2015, ISBN 978-80-244-4509-0
2011	J. Jáčová, J. Hřebíčková: <i>Laboratorní metody</i> . Univerzita Tomáše Bati ve Zlíně [online], 2011. [cit. 2019-06-05]. Dostupné z WWW: < <a href="http://ads.fmk.utb.cz/Contexts/atelier/Documents/Laboratorn%C3%AD%20metody.pdf">http://ads.fmk.utb.cz/Contexts/atelier/Documents/Laboratorn%C3%AD%20metody.pdf</a> >

## OCENĚNÍ

---

2018	cena děkana LF UP za nejlepší studentskou vědeckou práci za rok 2017
2017	Best Poster Prize at MetaboMeeting 2017, Birmingham, UK
2001	cena Nadačního fondu Miroslava Jurečka za vynikající úroveň diplomové práce Chemicko-technologické fakulty Univerzity Pardubice

## KONFERENCE

---

2019	HPLC 2019, Miláno, Itálie, 16. – 18. 6. 2019, přednáška: „Diagnosing of organic acidurias and beta-oxidation defects by LC-MS/MS“
2019	34. pracovní dny dědičné metabolické poruchy, Olomouc, Česká republika, 29. 5. – 31. 5. 2019, přednáška: „Diagnostika organických acidurií a poruch $\beta$ -oxidace mastných kyselin pomocí LC-MS/MS
2019	Česká konference hmotnostní spektrometrie 2019, 27.3. – 29. 3. 2019, Olomouc, Česká republika, přednáška: „Diagnosing of organic acidurias and beta-oxidation defects by LC-MS/MS“
2018	MetaboMeeting 2018, Nottingham, UK, 17. – 19. 12. 2018, poster: „Profiles of organic acids by liquid chromatography-mass spectrometry“
2018	MOVISS, Vorau, Rakousko, 9. – 12. 9. 2018, pasivní účast
2018	15th GCXGC Symposium, Riva del Garda, Itálie, 13. 5 – 18. 5. 2018, plenární přednáška „Bottlenecks of orthogonality evaluation“
2018	42nd International Symposium on Capillary Chromatography, Riva del Garda, Itálie, 13. 5 – 18. 5. 2018, poster „ „Urease immobilized on magnetic microparticles for GC/MS-based urinary metabolomics“
2017	Metabomeeting 2017, Birmingham, UK, 10. – 13. 12. 2017, poster „Urease immobilized on magnetic microparticles for GC/MS-based urinary metabolomics“
2017	MOVISS 2017, Vorau, Rakousko, 20. - 22. 9. 2017, pasivní účast
2017	41st International Symposium on Capillary Chromatography & 14th GCxGC Symposium, Texas, USA, 14. - 19. 5. 2017, přednáška „The impact of normalization on orthogonality evaluation“
2017	6. konference České společnosti pro hmotnostní spektrometrii, Olomouc, Česká republika, 29. - 31. 3. 2017, přednáška „Dimenzionalita vzorku v optimalizaci 2D separací“

- 2016 Annual Symposium of the Society for the Study of Inborn Errors of Metabolism - SSIEM 2016, Řím, Itálie, 6. – 9. 9. 2016, poster: „Urease immobilized on magnetic microparticles for urine GC/MS analysis“
- 2016 Advances in chromatography and electrophoresis & Chiranal 2016, Olomouc, Česká republika, 6. 6. – 9. 6. 2016, přednáška „Evaluation of orthogonality assessment tools in GC×GC“.
- 2016 ISCC - 40th ISCC and 13th GCxGC Symposium, Riva del Garda, Itálie, 29. 5. - 3. 6. 2016, vyžádaná přednáška „Evaluation of orthogonality assessment tools in GC×GC“.
- 2016 31. Pracovní dny dědičné metabolické poruchy, Bratislava, Slovenská republika, 25.5. – 27. 5. 2016, přednáška „Ureáza imobilizovaná na magnetických mikročásticích pro metabolomiku lidské moči“.
- 2016 5. konference České společnosti pro hmotnostní spektrometrii, České Budějovice, Česká republika, 13.-15. 4. 2016, přednáška „Ureáza imobilizovaná na magnetických mikročásticích pro metabolomiku lidské moči“.
- 2015 30. Pracovní dny Dědičné metabolické poruchy, Dolní Morava, Česká republika, 20. – 22. 5. 2015, poster „Stanovení galaktitolu v moči pomocí GC/MS“
- 2014 XXXIII. Regionální pracovní dny klinické biochemie, Karlova Studánka, česká republika, 3. – 5. 12. 2014, přednáška „Kontrola kvality v diagnostice metabolických chorob“
- 2014 38th International Symposium on Capillary Chromatography (ISCC) and 11th GCxGC Symposium, Riva del Garda, Itálie, 18. – 23. 5. 2014, poster „Evaluation of different column combinations in GCxGC-TOF MS analysis of organic acids in human urine“

# UNIVERZITA PALACKÉHO V OLOMOUCI

Přírodovědecká fakulta  
Katedra analytické chemie



## AUTOREFERÁT

k disertační práci

### Vícerozměrné separační techniky pro analýzu metabolomu lidské moči

Autor práce:

RNDr. Jaroslava Jáčová

Studijní program:

P1407 Chemie

Studijní obor:

Analytická chemie

Vedoucí práce:

prof. RNDr. Tomáš Adam, Ph.D.

Konzultanti:

doc. RNDr. David Friedecký, Ph.D.,

prof. Jean-Marie D. Dimandja, Ph.D.

Olomouc 2019

## Souhrn

Předkládaná disertační práce se zabývá instrumentálními analytickými metodami v metabolomice. Metabolomika je vědní disciplína umožňující detailní pohled na celkový profil nízkomolekulárních analytů (metabolitů) v biologických materiálech. Analytické nástroje proto musí čelit velké rozmanitosti a koncentračnímu rozsahu sledovaných analytů při vysoké komplexitě vzorků.

V první části je vyvinut nový přístup přípravy vzorku na základě imobilizované ureasy, který umožňuje maximální zachování metabolomu moči při eliminaci obvyklých artefaktů a interferentů. Nově vyvinutá metoda je porovnávána s relevantními již dříve publikovanými postupy (volnou ureasou, kyselou extrakcí do ethylacetátu) a vzorky ponechanými bez úpravy. Přístup založený na imobilizované urease je úspěšně aplikován na necílenou metabolomickou analýzu vzorků moči pacientů s dědičnou metabolickou poruchou (glutarovou acidurií typu I) a zdravých kontrol.

V rámci zaměření na vícerozměrné separační metody, zejména GC×GC, je představen nový způsob škálování dvojdímenzionálního separačního prostoru, WOSEL (z angl. whole separation space scaling) založený na analytické metodě. Při aplikaci WOSEL je separační prostor ohraničen mrtvým časem na jedné straně a koncem separačního záznamu na straně druhé. Oproti tomu u tradičních přístupů je separační prostor zcela nebo zčásti definován retenčním časem analytů (nejvíce a nejméně zadržovaným analytem daného vzorku). Oba dříve publikované přístupy jsou proto závislé na vzorku použitém k optimalizaci. Při škálování pomocí WOSEL lze díky tomu s výhodou použít jakýkoli vzorek podobného typu. Navíc nedochází ke zkreslení separačního prostoru (i po škálování odpovídá vzhledu chromatogramu).

Dále je pozornost věnována nalezení optimálního způsobu určení ortogonality dvojdímenzionálních chromatografických systémů. Je testováno celkem osm deskriptorů, které ortogonalitu počítají na základě odlišných principů. Žádný z deskriptorů ale není schopen ohodnotit homogenitu pokrytí (lokální ortogonalitu) a rozsah využitého separačního prostoru (globální ortogonalitu) najednou. Proto je pomocí geometrického průměru zkombinován nejlepší deskriptor lokální a globální ortogonality do nového deskriptoru nazývaného ASCA (z angl. AN and SC average), zahrnující oba aspekty ortogonality. Pomocí tohoto deskriptoru je nalezena optimální kombinace kolon pro GC×GC-MS analýzu metabolomu moči a plasmu.

## Summary

The thesis deals with instrumental analytical methods in metabolomics. Metabolomics is a scientific discipline that focuses on the low molecular weight analytes (metabolites) contained in biological materials. Therefore, analytical tools must face the great diversity and concentration range of the analytes analyzed and a high sample complexity.

In the first part of the thesis, a new approach to sample preparation based on immobilized urease is developed, which allows maximum retention of urine metabolome by eliminating common artefacts. This newly developed method is compared with methods already published (free urease, acidic extraction into ethyl acetate) and samples left untreated. The approach based on immobilized urease has been successfully applied to untargeted metabolomic analysis of urine samples from patients with inherited metabolic disorder – glutaric aciduria type I, and healthy controls.

Focused on multidimensional separation methods, especially GC × GC, a new method of scaling of a two-dimensional separation space, Whole Separation Space Scaling (WOSEL), is introduced. When using WOSEL, the separation space in both dimensions begins with the void time and ends with the last time point of data acquisition. Compared to WOSEL traditional scaling methods applies separation space delineated by the retention time of the most (and in one case least) retained sample analyte. Both previously published approaches are dependent on the sample used for optimization. In the case of WOSEL, any sample of a given (or similar) type of biological material can be used. Moreover, the space in which orthogonality is calculated corresponds to the perceived quality of the chromatogram (no separation space distortion occurs).

Furthermore, attention is paid to finding the optimal way of determining orthogonality of two-dimensional chromatographic systems. A total of eight descriptors are tested that calculate orthogonality based on different principles. However, none of the descriptors is able to evaluate homogeneity of coverage (local orthogonality) and the extent of separation space used (global orthogonality) at once. Therefore, using the geometric mean, the best local and global orthogonality descriptor is combined into a new descriptor called ASCA (i.e. AN and SC average), encompassing both aspects of orthogonality. Using this descriptor, an optimal column combination for GC × GC-MS urine and plasma metabolome analysis is evaluated.



# OBSAH

1	Úvod .....	5
2	Cíle práce .....	7
3	Imobilizovaná ureasa pro GC-MS metabolomiku moči .....	8
3.1	Imobilizovaná ureasa .....	8
3.2	Vzorky .....	9
3.3	Metody .....	10
3.4	Výsledky a diskuse .....	10
3.5	Závěr .....	15
4	WOSEL jako nový způsob škálování 2D separačního prostoru .....	16
4.1	Metody .....	18
4.2	Výsledky a diskuse .....	18
4.3	Závěr .....	20
5	Význam dimenzionality vzorku pro výpočet ortogonalitu 2D separací .....	21
5.1	Metody .....	22
5.2	Výsledky a diskuse .....	22
5.3	Závěr .....	23
6	Nalezení nejvíce ortogonální kombinace kolon pro GC×GC analýzu metabolomu moči ...	25
6.1	Metody .....	25
6.2	Výsledky a diskuse .....	25
6.3	Závěr .....	27
7	Seznam použité literatury .....	28
	Konference .....	34

# 1 Úvod

Metabolom je heterogenní skupinou chemických látek, jejichž společným znakem je pouze nízká molekulová hmotnost (pod 1500 Da) a výskyt v daném biologickém materiálu. Metabolom člověka obsahuje látky 356 různých chemických tříd. V moči se vyskytují především látky hydrofilní, zatímco v krvi převažují lipidy a mastné kyseliny, které v ledvinách do moči téměř nepřecházejí. Zastoupení hydrofilních látek v moči a séru je téměř identické, přesto lze v moči detekovat přibližně o 500 látek více než v krvi. Příčinou je koncentrační schopnost ledvin, které při zpětném vstřebávání vody nepřímou zvýší koncentraci metabolitů v moči až o tři řády. To, společně se snadným (neinvazivním) odběrem, činí moč vyhledávaným diagnostickým materiálem.

Studiem metabolomu se zabývá věda nazývaná metabolomika. Prostřednictvím specifických metabolických profilů studovaných systémů nabízí lepší porozumění jejich biochemickým a patobiochemickým procesům, odhaluje mechanismy onemocnění, identifikuje nové biomarkery a umožňuje personalizovanou léčbu onemocnění či sledování účinků terapie. Metabolomika hledá rozdíly ve složení komplexního profilu nízkomolekulárních látek mezi alespoň dvěma skupinami vzorků (nejčastěji mezi skupinou pacientů a zdravých kontrol).

Metabolomika řeší dva základní typy problémů: buď sleduje a porovnává abundance předem vybraných, známých metabolitů u sledovaných metabolických drah, tzv. cílená metabolomika, nebo ve vzorcích vyhledává a zaznamenává všechny látky vybraného rozsahu molekulových hmotností bez prvotní znalosti jejich identifikace, tzv. necílená metabolomika. Cílená metabolomika tak přináší informace o využití substrátů a produktů metabolických reakcí, fungování metabolických drah sledovaného biologického systému, umožňuje popsat a hlouběji porozumět jeho biochemii a patobiochemii. Oproti tomu necílená metabolomika vzhledem k obvykle vyššímu pokrytí metabolomu umožňuje formulovat nové hypotézy, odhalit nové biochemické dráhy a biomarkery onemocnění [1]. Oba přístupy poskytují z části komplementární informace, a proto se často provádějí současně.

Velká chemická různorodost cílových látek klade vysoké nároky na analytické nástroje používané v metabolomice. Důležitá je dostatečná citlivost ( $< \mu\text{mol.l}^{-1}$ ), velký dynamický rozsah měření (rozdíly v koncentracích mezi jednotlivými analyty jsou až 12 řádů), vysoká frakcionační kapacita (některé ze sledovaných metabolitů jsou si strukturně velmi podobné) a dostatečná stabilita systému (často se měří série až tisíců vzorků).

V současné době neexistuje analytický systém, který by byl sám schopen měřit celý metabolom. Proto se metabolomické studie zaměřují buď na určitou část metabolomu (např. organické kyseliny) nebo se použije kombinace několika technik. Pro akvizici dat se většinou používá nukleární magnetická rezonance (NMR) nebo hmotnostní spektrometrie (MS). Zatímco NMR vyniká vysokou reprodukovatelností výsledků a minimální přípravou vzorků, MS citlivostí. Frakcionační kapacita MS je však velmi nízká, proto bývá spojována s některou ze separačních technik: nejčastěji s kapalinovou chromatografií, méně pak s plynovou chromatografií a zřídka i s kapilární elektroforézou. Nevýhodou GC je zdouhavá příprava vzorku a vyšší detekční limit (řádově v  $\mu\text{mol.l}^{-1}$ ), u LC se jedná o nízkou reprodukovatelnost retenčních časů, nižší životnost kolon a separační účinnost.

Při analýze komplexních vzorků se ve spojení s MS s výhodou používají i dvojrozměrné chromatografické systémy, ve kterých separace v každé z dimenzí probíhá na základě jiného mechanismu, čímž se mnohonásobně navýší frakcionační kapacita. Příkladem takového systému je dvojdimenzionální plynový chromatograf (GC×GC), který obsahuje dvě sériově uspořádané kolony o různé selektivitě spojené prostřednictvím modulátoru. To, zda takový systém představuje benefit pro analýzu konkrétního typu vzorku, se ověřuje mnoha způsoby. Mezi nejznámější patří stanovení míry nárůstu píkové kapacity a určení ortogonalit. V této práci je využito ortogonálního přístupu.

Datové soubory vyhodnocované v metabolomických studiích jsou komplexní, obsahující stovky proměnných (abundance jednotlivých metabolitů) od desítek až stovek pozorování (vyšetřovaných subjektů - vzorků). Vlastnímu vyhodnocení dat předchází jejich předběžná analýza a úprava. Patří do ní např. nastavení základní linie hladiny šumu, správné zarovnání píků analytů (angl. „peak alignment“), nahrazení chybějících hodnot, nastavení parametrů integrace píků, korekce na nestabilitu analytického systému, škálování (převedení abundancí proměnných na stejné měřítko), transformace (data získají normální rozdělení) centrování (nastavení střední hodnoty proměnné na hodnotu 0).

Následuje průzkumová analýza dat, která zahrnuje testování datového souboru na normalitu, korelace, homogenitu a určení odlehlých hodnot. Tento krok je nezbytný pro správný výběr statistických testů. Pokud mají data normální rozdělení, lze použít parametrické testy (např. t-test), v opačném případě je nutné použít neparametrické testování (např. Mann-Whitneyův U-test). Vlastní analýza dat metabolomického experimentu je pak provedena vícerozměrnými statistickými metodami (např. pomocí analýzy hlavních komponent, PCA).

## 2 Cíle práce

Cíle disertační práce lze shrnout do následujících bodů:

- I. Zavést přístup přípravy vzorku pro analýzu metabolomu pomocí GC-MS založený na urease imobilizované na magnetických mikročástečkách a porovnat jej s dalšími metodami přípravy vzorku (kyselou extrakcí do ethylacetátu, volnou ureasou, vzorkem bez úpravy).
- II. Vyvinout způsob škálování 2D separačního prostoru, který by co nejuvěrněji odrážel využitelný separační prostor a byl nezávislý na vzorku, použitém pro optimalizaci.
- III. Najít nejvhodnější způsob určení ortogonality pro nízkodimenzionální a vysokodimenzionální vzorky.
- IV. Pomocí ortogonality určit nejvhodnější kombinaci kolon pro GC×GC analýzu moči a porovnat ji s kombinací pro plasmatické vzorky.

### 3 Imobilizovaná ureasa pro GC-MS metabolomiku moči

Močovina komplikuje GC-MS analýzu, protože jako konečný produkt metabolismu bílkovin je v moči přítomna ve vysoké koncentraci. Při přípravě vzorku pak dochází ke konkurenci mezi analyty a močovinou při derivatizaci, a tak je řada analytů derivatizována jen částečně anebo vůbec. V separaci jsou místo jednoho derivatizačního produktu daného analytu přítomny dva nebo více, což má zásadní vliv na limit detekce těchto látek. Volné vodíkové atomy nederivatizovaných skupin se ireverzibilně váží na povrch stacionární fáze, čímž se zhoršuje tvar píku analytu i životnost kolony [2]. Samotná močovina v koloně působí jako separační interference: vyvolává přetížení kolony, její tři píky (nederivatizovaná močovina, močovina derivatizovaná dvěma a třemi trimethylsilylovými (TMS) skupinami) maskují řadu analytů a tím znesnadňují jejich detekci či kvantifikaci. Proto je nezbytné obsah močoviny před GC-MS analýzou moči snížit nebo ji zcela odstranit.

Tradičním a dosud rutinně používaným přístupem pro přípravu vzorku moči pro GC-MS analýzu je kyselá extrakce do ethylacetátu [3]. Močovina je při pH 1-2 ionizována, a tak do ethylacetátu přechází jen v omezené míře. Zároveň však dochází k ionizaci a tím i snížené extrakci celé řady dalších metabolitů moči (např. aminokyseliny, hydroxykyseliny, acylglyciny, acylkarnitiny, puriny, pyrimidiny). Metabolom moči se tak výrazně změní. Navíc účinnost extrakce je proměnlivá, což má za následek vysokou analytickou variabilitu.

Alternativním přístupem je aplikace ureasy [2,4–7], enzymu izolovaného z kanaválie mečovité (*Canavalia ensiformis*, někdy též nazývané bob koňský nebo fazole gotani). Tento enzym ze třídy hydrolas (E.C. 3.5.1.5) rozkládá močovinu na oxid uhličitý a amoniak [8,9]. Je vysoce specifický, z dalších substrátů hydrolyzuje pouze hydroxymočovinu (lék pro léčbu klinické myeloidní leukémie [10]) na  $\text{HCO}_3^-$ ,  $\text{NH}_4^+$  a  $\text{NH}_2\text{OH}$  [9]. Tento přístup má tři hlavní nevýhody: je poměrně pracný [2,4], zanáší do analýzy řadu artefaktů (převážně při odstraňování ureasy deproteinací) [4–6,11] a vyvolává změny hladin některých metabolitů [5,12]. V této práci se nově využívá ureasy imobilizované na magnetický nosič. Testuje se vliv imobilizace na další nakládání s ureasou, množství artefaktů a na metabolom moči jako celek a porovnává se imobilizovaná ureasa s ostatními přístupy přípravy vzorku.

#### 3.1 Imobilizovaná ureasa

Ureasa imobilizovaná na magnetické mikročástice byla vyvinuta na Katedře biochemie Přírodovědecké fakulty Univerzity Palackého v Olomouci [13].

## 3.2 Vzorky

V experimentu je použito dvou typů vzorků: kontrolního materiálu na bázi lidské moči (COA, z angl. „control organic acids“), do kterého je výrobcem přidána známá koncentrace 23 metabolitů, a moči pacientů. Pro analýzu je použit objem vzorku odpovídající koncentraci kreatininu 1 mmol.l<sup>-1</sup>.

Vzorek COA je slouží ke stanovení vlivu přípravy vzorku (zejména nově vyvinutého přístupu založeného na imobilizované urease) na metabolom moči. COA je připraven 4 různými metodami (vždy 6 replikátů + blank pro každou metodu):

- 1) Kyselá extrakce do ethylacetátu (EX): ke vzorku COA je přidán interní standard (IS, 4-fenylbutyrát) a O-ethylhydroxylamin hydrochlorid (derivatizace I). Po 30 min stání při pokojové teplotě je vzorek okyselen 3 kapkami zředěné HCl (1:1), iontová síla upravena pomocí NaCl. Následuje třístupňová extrakce do ethylacetátu. Extrakt je poté vysušen bezvodým Na<sub>2</sub>SO<sub>4</sub>, dekantován do čisté zkumavky a odfoukán proudem dusíku při 30 °C. Odparek je 2x vymyt methanolem, a pak ještě jednou acetonem, znovu odfoukán proudem dusíku při 30 °C. Nakonec je přidána derivatizační směs (BSTFA:pyridin:TMCS, 25:25:1 (v/v/v), derivatizace II). Vialka je uzavřena pod dusíkem, 20 min udržována při 70 °C.
- 2) Volná ureasa (FUT): COA je smíchán s roztokem ureasy, 40 min promícháván při 37 °C, poté je enzym precipitován methanolem (30 min, -80 °C) a centrifugován. Supernatant je přenesen do skleněné vialky obsahující IS a zlyofilizován. Následně je vzorek derivatizován ve dvou krocích: ethoximace (60 min při 50 °C), trimethylsilylace (20 min při 70 °C).
- 3) Imobilizovaná ureasa (IUT): COA je smíchán se suspenzí částic s imobilizovanou ureasou, 40 min promícháván na míchačce při 37 °C. Magnetické mikročástice jsou potom odděleny pomocí magnetu, po přidání IS je vzorek zlyofilizován. Následně je vzorek derivatizován ve dvou krocích: ethoximace (60 min při 50 °C), trimethylsilylace (20 min při 70 °C).
- 4) Vzorky bez úpravy (NT): ke COA je přidán IS. Vzorek je zlyofilizován a následně derivatizován ve dvou krocích: ethoximace (60 min při 50 °C), trimethylsilylace (20 min při 70 °C).

Vzorky moči pro metabolomický experiment jsou připraveny pomocí imobilizované ureasy (10 vzorků pacientů s glutarovou acidurií typ I, 10 vzorků kontrol a jeden blank). Vzorek kontroly kvality (QC) je připraven smícháním 10 µl z každého vzorku (vyjma blanku). Vzorky jsou analyzovány pomocí GC-MS, a to v náhodném pořadí. QC je zařazeno jako každý 6. vzorek.

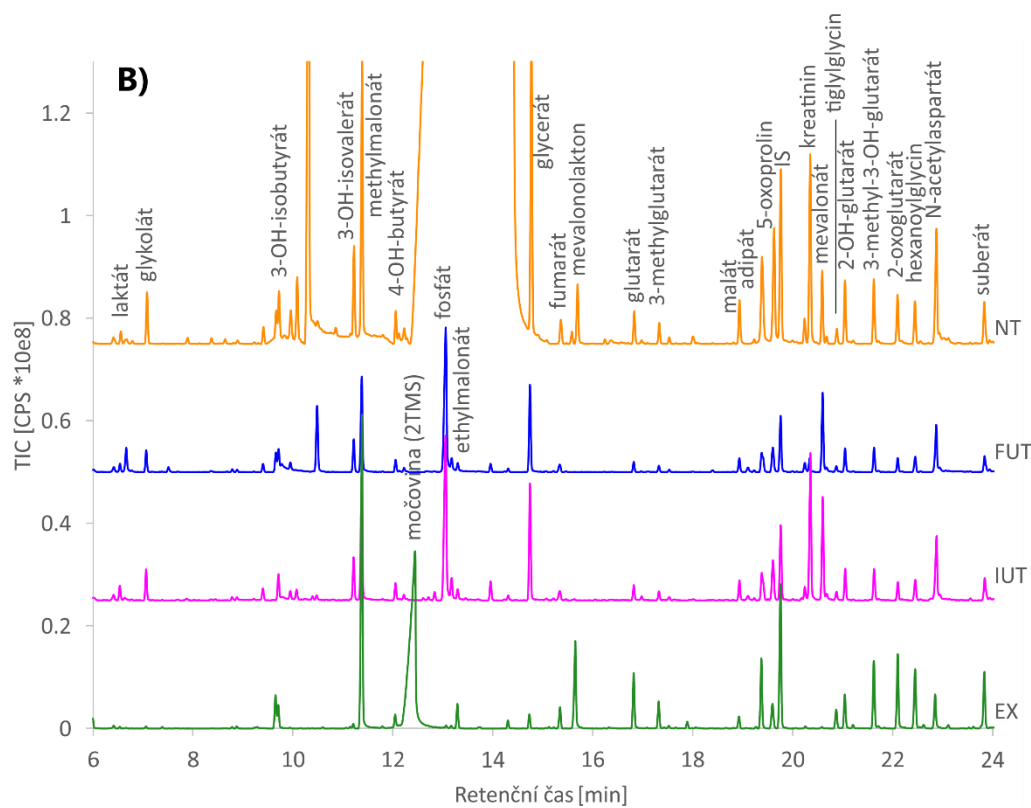
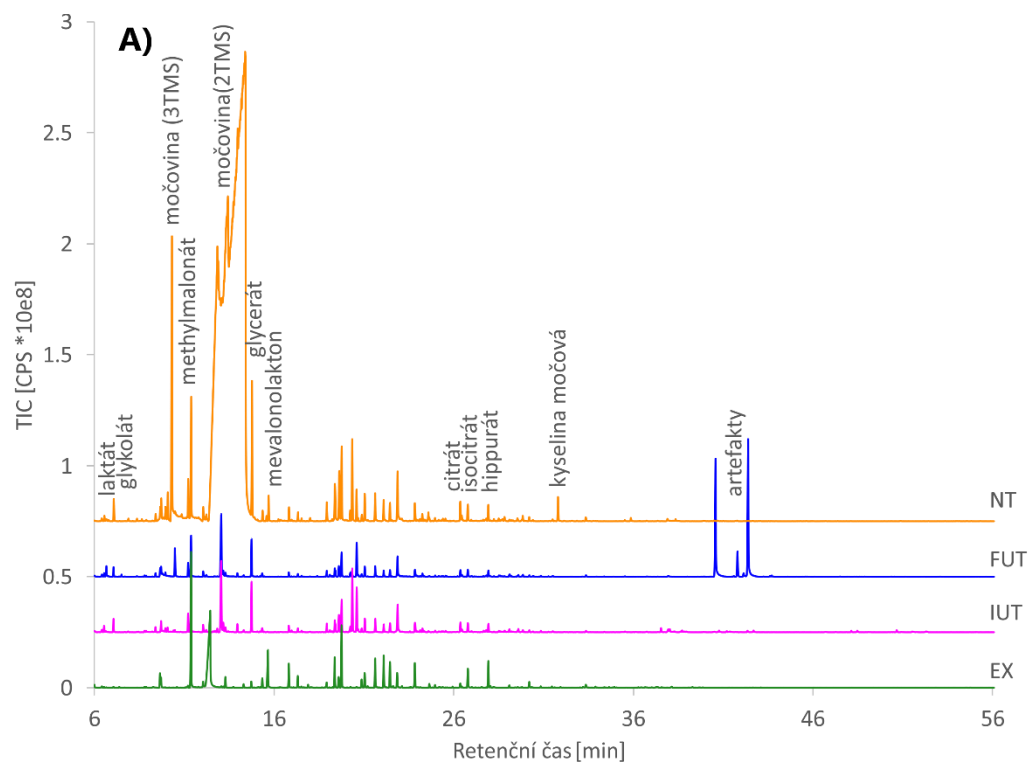
### 3.3 Metody

Vzorky jsou analyzovány na plynovém chromatografu spojeném s TOF MS (Pegasus4, LECO, St. Joseph, Michigan, USA), vybaveném kolonou ZB-5MS (30 m; 0,25 mm; 0,25  $\mu$ m; Phenomenex, Torrance, California, USA). Jako nosný plyn je použito helium v režimu konstantního průtoku (1 ml.min<sup>-1</sup>). Podíl vzorku (1  $\mu$ l) je zaváděn do proudu nosného plynu metodou horké jehly při 250 °C, v režimu s děličem toku s dělicím poměrem 1:10. Je využito analýzy s teplotním programem. K akvizici dat slouží hmotnostní spektrometr s analyzátorem doby letu (TOF) a elektronovou ionizací (EI). Datový soubor je procesován v R-software (verze 3.3.0, <https://www.r-project.org>) pomocí softwarového balíčku vyvinutého v naší laboratoři [14], za použití *clr* transformace a XCMS. Ke zobrazení dat slouží zejména PCA.

### 3.4 Výsledky a diskuse

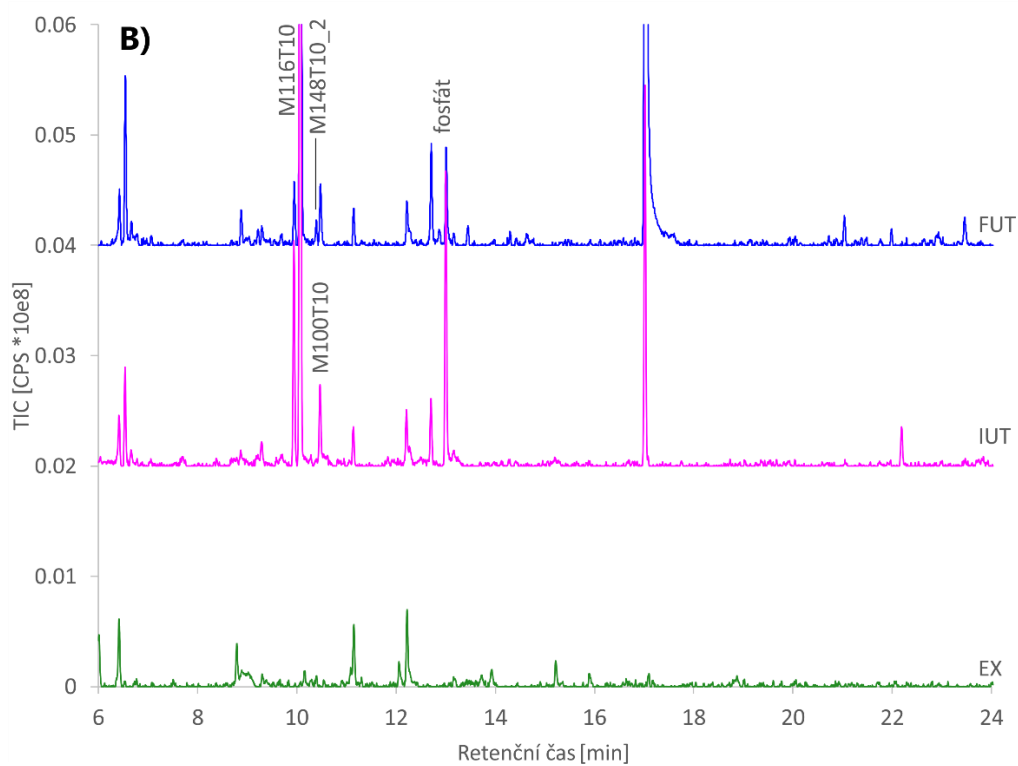
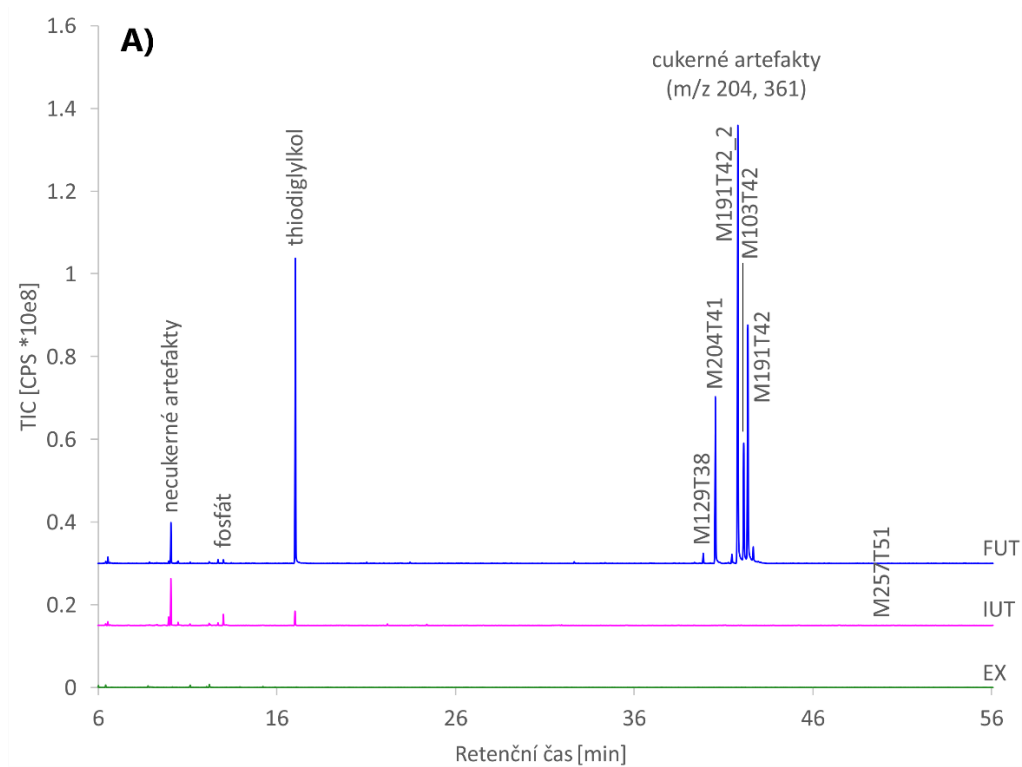
Testované metody přípravy vzorku se liší v pracnosti i časových nárocích. Nejpracnější ze všech je EX, který se skládá z 11 různých kroků a 6 h nepřetržité práce. Vzorek je však k analýze připraven týž den, protože nevyžaduje lyofilizaci. Většinou je ale obtížné připravit více než 12 vzorků najednou. Při použití FUT lze vzorek připravit v 8 různých krocích, vyžaduje 3 h práce a lyofilizaci přes noc. Oproti tomu IUT zahrnuje pouze z 5 kroků. Protože nepotřebuje deproteinaci, k přípravě jsou nutné pouze 2 h práce a lyofilizace přes noc. Oba ureasové protokoly umožňují přípravu většího množství vzorků. Jediným omezením je kapacita rotátoru.

Chromatografický vzor GC-MS analýzy IUT a FUT je obdobný (Obr. 1). Hlavním rozdílem mezi analýzami je řada unikátních píků na konci analýz FUT. Tyto píky jsou přítomny i v analýzách blanků (Obr. 2), proto se lze domnívat, že se jedná o artefakty. Oproti tomu u EX je přítomno nejméně píků, navíc s naprosto odlišnou odezvou. NT vzorky mají nejvyšší odezvy v celkovém iontovém proudu (TIC) vzorků, obzvláště pak v retenčním okně močoviny (mezi 12. a 15. min). Ohromný pík močoviny také způsobuje přetížení kolony, a proto i posun retenčních časů látek, které jsou v koloně zadržovány více než močovina (např. glycerát, fumarát, mevalonolakton). Močovina také zhoršuje identifikaci koeluuujících látek (např. pro ethylmalonát v NT vzorcích byla shoda s NIST pouze okolo 300, což není pro identifikaci látky dostačující).



Obr. 1: Analýza COA připraveného pomocí jednotlivých způsobů přípravy vzorku: A) celkový chromatografický vzor, B) detail.

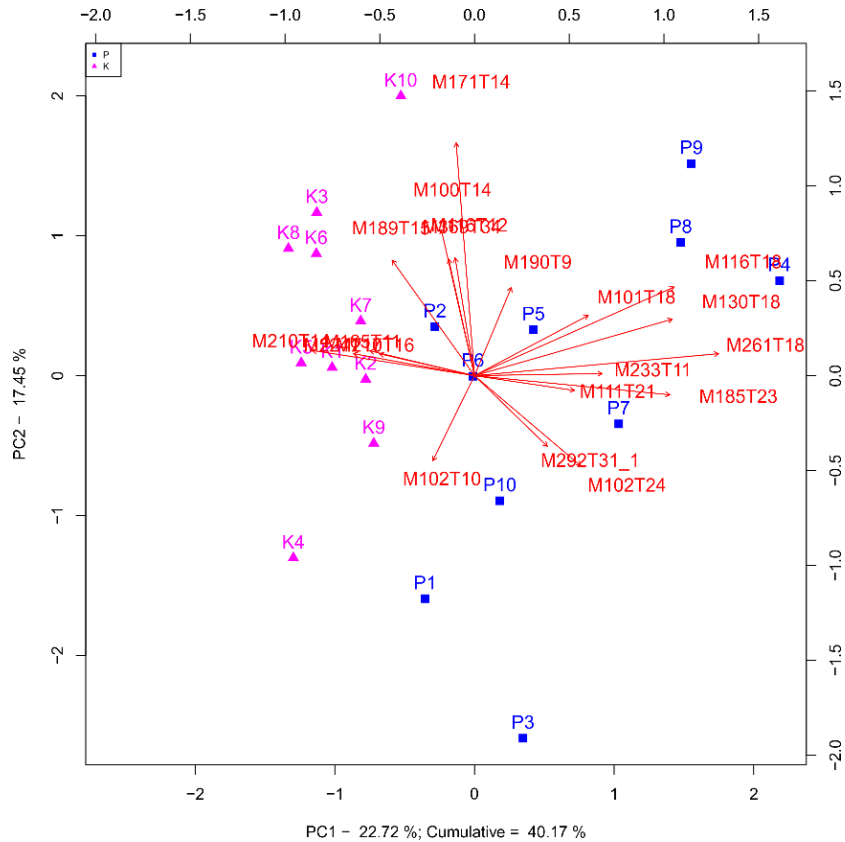




Obr. 2: Analýza blanku připraveného pomocí jednotlivých způsobů přípravy vzorku: A) celkový chromatografický vzor, B) detail.



onemocnění způsobené deficitem glutaryl-CoA-dehydrogenasy (EC 1.3.99.7), díky kterému nejsou postižení jedinci schopni odbourávat lysin, hydroxylysin a tryptofan. Většinou se projevuje mozkovou atrofií způsobující změny v motorice a hybnosti, které jsou při zátěži organismu výraznější, často také i mentální retardaci [15]. V moči jsou přítomny zvýšené hladiny 3-hydroxyglutarátu a glutarátu [16].



Obr. 4: Dvojný graf PCA necílené GC-MS analýzy močí pacientů s glutarovou acidurií I (P1-P10) a zdravých kontrol (K1-K10), připravených pomocí imobilizované ureasy.

Vzorky jsou připraveny pomocí IUT, data jsou zpracována necíleným metabolickým přístupem. Ve vzorcích je nalezeno celkem 221 píků metabolitů. Pomocí PCA (Obr. 4) se vzorky dělí na 2 skupiny. V první skupině jsou zastoupeny pouze zdravé kontroly, ve druhé pacienti s GAI. U pacientů jsou v porovnání s kontrolami přítomny signifikantně zvýšené hladiny 3-hydroxyglutarátu (M185T23, RI 1575, shoda s knihovnou 968), tartarátu (M102T24, RI 1636, shoda s knihovnou 917), glutarátu (M261T18, RI 1403, shoda s knihovnou 880) a neznámé látky M111T13 (RI 1213). Snížené hladiny jsou pozorovány pro M210T14 (RI 1240), M224T17 (RI 1362), M210T16 (RI 1317), M210T17 (RI 1350) a M111T21 (RI 1507). Glutarát a 3-hydroxyglutarát jsou známými biomarkery onemocnění, tartarát je součástí léků podávaných pacientům s GAI. Neznámé látky M210T14, M224T17, M210T16, M210T17 a M111T21 jsou komponenty

moči zdravých jedinců, jež se nepodařilo identifikovat pomocí NIST ani GOLM databáze. Jejich identifikace bude předmětem další práce.

### 3.5 Závěr

Ureasa imobilizovaná na magnetických částicích je novou technikou přípravy vzorku moči pro GC-MS analýzu. Ve srovnání s ostatními přístupnými přípravami (kyselá extrakce do ethylacetátu, volná ureasa, vzorek bez úpravy) vyniká snadností přípravy i reprodukovatelností výsledků (pro 23 organických kyselin a acylglycinů obsažených v referenční moči činí průměrné CV 2,5 %).

Chromatogramy vzorků připravených pomocí imobilizované ureasy jsou podobné přístupu za použití volné ureasy, v analýze je však méně artefaktů (jen 3 z původních 10). Je tomu tak nejspíš proto, že imobilizovaný enzym lze před aplikací vzorku snadno a opakovaně promýt. Vzorky připravené extrakcí poskytují zcela odlišné chromatografické záznamy. Ve srovnání s oběma ureasovými protokoly obsahují zbytky močovininy a také odlišné hladiny metabolitů (především kreatininu a glycerátu). Chromatografický systém je přetížen přítomností vysoké koncentrace močovininy, která vede ke změnám retence i odezvy metabolitů, zejména ethylmalonátu.

Prostřednictvím metabolomického experimentu se vzorky pacientů s glutarovou acidurií I je ověřeno, že imobilizovanou ureasu lze použít i pro necílenou GC-MS metabolomiku. Chromatografická data je možné zpracovávat běžným způsobem. Hlavními diskriminujícími metabolity jsou biomarkery GAI, a tak lze usuzovat, že aplikace imobilizované ureasy neovlivňuje metabolom moči.

## 4 WOSEL jako nový způsob škálování 2D separačního prostoru

Dvojměrné chromatografické systémy jsou tvořeny dvěma separačními dimenzemi, které se liší ve svém rozsahu [17–19]. Např. v GC×GC může separace v první dimenzi trvat 3000 s, zatímco v druhé pouze 4 s, a tak se separace v dosažená v druhé dimenzi může jevit zanedbatelnou. Pokud by to byla pravda, pak by ale připojení další separační dimenze bylo zbytečnou komplikací a ne přínosem.

Proto, aby se kladl stejný důraz na separaci dosaženou v první a druhé dimenzi, se retenční data před vyhodnocením efektivity dané separace škálují. Škálování má ale ještě další dva důsledky: tvoří normalizovaný retenční prostor (a tak je někdy označováno též jako normalizace) a odstraňuje nevyužitý či nevyužitelný separační prostor.

V pracích zabývajících se optimalizací dvojdímných systémů pomocí ortogonalit se škálování 2D separačního prostoru provádí dvěma různými způsoby, které lze podle hranic, jež retenční prostor vymezují, označit jako Void-max [17] a Min-max [20]. Oba způsoby škálování jsou pro první dimenzi znázorněny na Obr. 5. U staršího z nich, Void-max, je normalizovaný separační prostor je pro vymezen mrtvým časem dané dimenze a retenčním časem nejvíce zadržované složky vzorku použitého pro optimalizaci (Rov. 1).

$$X(A) = \frac{t_r(A) - t_0}{t_r(max) - t_0} \quad (\text{Rov. 1}),$$

kde  $X(A)$  je normalizovaný retenční čas látky A,  $t_r(A)$  je retenční čas látky a v dané dimenzi,  $t_0$  je mrtvý čas dané dimenze,  $t_r(max)$  je maximální retenční čas, kterého bylo v dané dimenzi dosaženo.

U Min-max algoritmu se retenční časy první a druhé dimenze škálují podle retenčního času nejméně a nejvíce zadržované látky (Rov. 2).

$$X(A) = \frac{t_r(A) - t_r(min)}{t_r(max) - t_r(min)} \quad (\text{Rov. 2}),$$

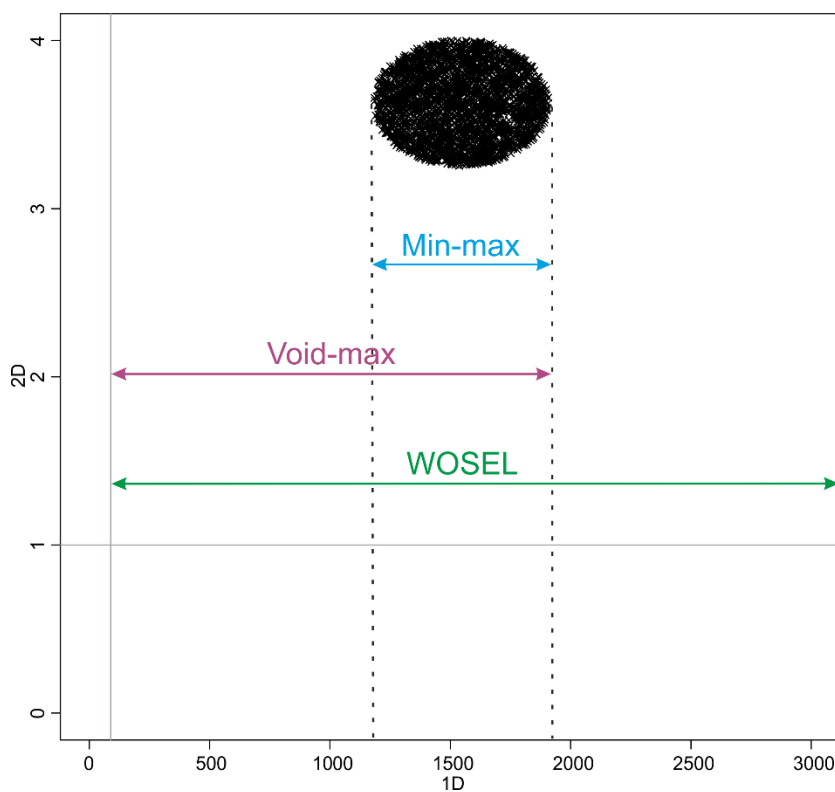
kde  $X(A)$  je normalizovaný retenční čas látky A,  $t_r(A)$  je retenční čas látky A v dané dimenzi,  $t_r(min)$  je retenční čas nejméně zadržovaného analytu,  $t_r(max)$  je retenční čas nejvíce zadržovaného analytu.

U obou výše uvedených přístupů je škálování závislé na vzorku použitém k optimalizaci. Jiný vzorek (obsahující jiné analyty) vymezí normalizovaný retenční prostor (nesrovnatelně) jinak. V případě cílené metabolomiky to nemusí představovat problém, protože je možné opakovaně namíchat stejnou směs standardů. V případě necílené metabolomiky, kdy všechny analyty přítomné ve vzorku nejsou předem známy, a navíc se přidává i vliv matrice, představuje takovéto škálování značnou komplikaci. Většinou se řeší tak, že se systém optimalizuje přímo pomocí daného biologického vzorku (většinou směsná moč či plasma). Jedná se o neopakovatelný vzorek, jehož množství je omezeno.

Cílem této části disertační práce je vyvinout škálování 2D prostoru, které by bylo nezávislé na konkrétním vzorku použitým k optimalizaci a co nejlépe odráželo původní separační prostor, a porovnat je pomocí syntetických (generovaných *in silico*) a reálných GC×GC chromatogramů se staršími škálovacími algoritmy. Nový přístup škáluje 2D separační prostor v souladu s analytickou metodou, tedy nezávisle na vzorku. Separační prostor je u něj vymezen mrtvým časem dané dimenze a posledním okamžikem sběru dat v dané dimenzi (u druhé dimenze je tento okamžik shodný s modulační periodou) (Rov. 3, Obr. 5). K výpočtu ortogonality využívá celého separačního prostoru, proto se postup škálování nazývá WOSEL (z angl. „WhOle SEparation space scaLing“).

$$X(A) = \frac{t_r(A) - t_0}{t_r(\text{end}) - t_0} \quad (\text{Rov. 3}),$$

kde  $X(A)$  je normalizovaný retenční čas látky  $A$ ,  $t_r(A)$  je retenční čas látky  $a$  v dané dimenzi,  $t_0$  je mrtvý čas dané dimenze,  $t_r(\text{end})$  je konec sběru dat v dané dimenzi.



Obr. 5: Různé způsoby škálování 2D separačního prostoru pro první dimenzi. Černý shluk představuje skupinu píků. Na osách je uveden retenční čas v sekundách.

## 4.1 Metody

Vliv typu škálování na výslednou ortogonalitu systému je ověřen pomocí syntetických a reálných GC×GC chromatogramů vytvořených v R-software (v. 3.3.0) tak, že tvoří oválný shluk obsahující 2500 datových bodů uvnitř separačního prostoru, který má v první dimenzi rozsah 89-3000 s (89 s je mrtvý čas první dimenze), ve druhé 1-4 s (1 s je mrtvý čas druhé dimenze). Body uvnitř shluku mají rovnoměrné rozdělení.

Ortogonalita je vypočítána několika různými přístupy vybranými tak, aby z každé skupiny přístupů o stejném principu byl vybrán právě jeden: Pearsonova korelace (PC) [21], podmíněná entropie (CE) [22], konvexní obálka (CH) [23], modelování (MA) [24], geometrické pokrytí separačního povrchu (SC) [20], metoda nejbližšího souseda (NN) [25] a hvězdičkové rovnice (AE) [26]. Retenční koordináty piků jsou před výpočtem ortogonalit škálovány třemi různými způsoby, uvedenými jako Rov. 1-3. Všechny výpočty jsou provedeny v R-software.

## 4.2 Výsledky a diskuse

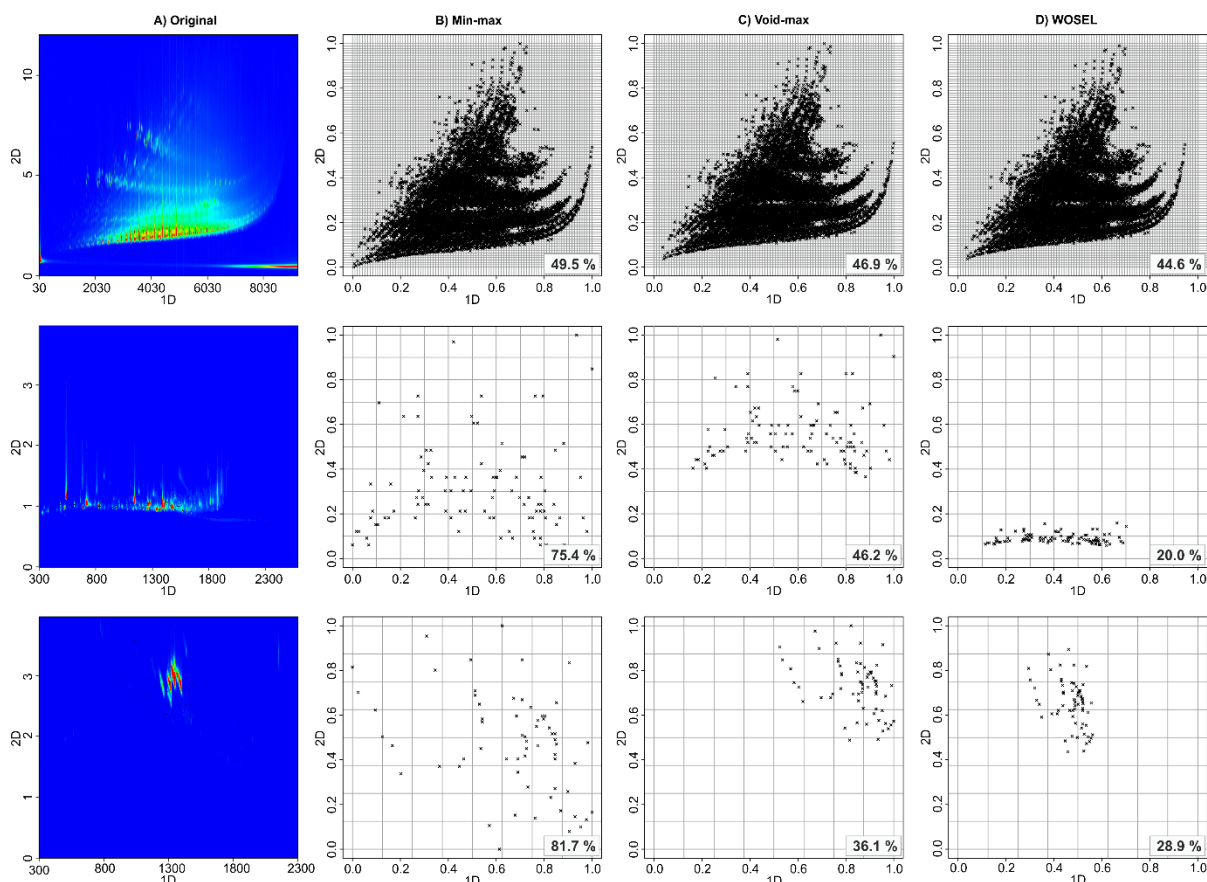
Při vyhodnocování optimálních dvojdimenčních separací dodnes hraje velkou roli vzhled chromatogramu (tzv. „percieved quality of chromatogram“) [27]. Jako dobrá separace bývá označována taková, ve které je využita velká část separačního prostoru a data tento prostor rovnoměrně pokrývají. Proto jsou výsledky prezentovány nejen v podobě číselných výsledků (Tab. 1, Obr. 6), ale je zde uveden vzhled výchozího chromatogramu (označený jako „Original“) a také příslušných normalizovaných retenčních prostorů, které vznikají z výchozího aplikací jednotlivých přístupů škálování: Min-Max (Rov. 2), Void-Max (Rov. 1) a WOSEL (Rov. 3).

Z Tab. 1 vyplývá, že zvolený způsob škálování ovlivňuje výsledky všech přístupů k výpočtu ortogonalit, s výjimkou PC, která si data v prvním kroku výpočtu sama škáluje (od každé hodnoty odečte průměr a pak ji vydělí směrodatnou odchylkou), a tak je nezávislá na použitém škálování [28]. Typ škálování má zásadní vliv na výslednou ortogonalitu zejména u vzorků, jejichž složky mají v 2D chromatogramu malou disperzi. V těchto případech starší způsoby škálování vedou k nadhodnocování ortogonalit chromatogramů, protože do normalizovaného retenčního prostoru převádí jen velmi malou část prostoru původního. Oproti tomu při použití WOSEL hodnota odpovídá vzhledu chromatogramu (Tab. 1, 78 vs. 9 %). Příkladem je i chromatogram plasmu uvedený na Obr. 6 (82 vs. 29 %). Naopak u vzorků, které vyplňují většinu separačního prostoru (Tab. 1), jako tomu je také u vzorku nafty na Obr. 6, má typ škálování na ortogonalitu vliv jen velmi malý.

Tab. 1: Průměrné hodnoty ortogonality jednotlivých typů modelů s 2500 body získané různými přístupy výpočtu ortogonality  $\pm 2SD$  pro 50 modelů stejného typu. Hodnoty jsou uvedeny v %, s výjimkou AN a HN, jejichž výsledkem jsou průměrné nejmenší vzdálenosti mezi nejbližšími sousedy.

Typ modelu	Způsob výpočtu	Min-Max	Void-Max	WOSEL
	SC	$88,2 \pm 1,8$	$87,8 \pm 1,6$	$87,7 \pm 1,6$
	MA	$88,2 \pm 1,8$	$87,8 \pm 1,6$	$87,7 \pm 1,6$
	CE	$85,0 \pm 0,7$	$85,0 \pm 0,6$	$85,1 \pm 0,3$
	PC	$98,7 \pm 1,9$	$98,7 \pm 1,9$	$98,7 \pm 1,9$
	AE	$79,4 \pm 2,3$	$78,7 \pm 2,2$	$78,3 \pm 2,2$
	CH	$78,2 \pm 0,7$	$77,6 \pm 0,6$	$77,1 \pm 0,2$
	AN	$0,0117 \pm 0,0002$	$0,0116 \pm 0,0002$	$0,0116 \pm 0,0002$
	HN	$0,0078 \pm 0,0006$	$0,0078 \pm 0,0006$	$0,0078 \pm 0,0006$
	SC	$87,9 \pm 1,7$	$32,0 \pm 0,5$	$8,7 \pm 0,2$
	MA	$87,9 \pm 1,7$	$32,0 \pm 0,5$	$8,7 \pm 0,2$
	CE	$85,0 \pm 0,6$	$91,6 \pm 0,7$	$95,3 \pm 0,6$
	PC	$98,8 \pm 1,9$	$98,8 \pm 1,9$	$98,8 \pm 1,9$
	AE	$79,1 \pm 2,2$	$28,7 \pm 1,0$	$4,9 \pm 0,2$
	CH	$78,2 \pm 0,7$	$19,4 \pm 0,1$	$4,8 \pm 0,0$
	AN	$0,0116 \pm 0,0002$	$0,0058 \pm 0,0001$	$0,0029 \pm 0,0000$
	HN	$0,0078 \pm 0,0006$	$0,0039 \pm 0,0002$	$0,0019 \pm 0,0002$
	SC	$87,9 \pm 1,7$	$8,6 \pm 0,2$	$8,6 \pm 0,2$
	MA	$87,9 \pm 1,7$	$8,6 \pm 0,2$	$8,6 \pm 0,2$
	CE	$85,0 \pm 0,6$	$95,3 \pm 0,6$	$95,3 \pm 0,6$
	PC	$98,8 \pm 1,9$	$98,8 \pm 1,9$	$98,8 \pm 1,9$
	AE	$79,1 \pm 2,2$	$4,9 \pm 0,2$	$4,9 \pm 0,2$
	CH	$78,2 \pm 0,7$	$4,8 \pm 0,0$	$4,8 \pm 0,0$
	AN	$0,0116 \pm 0,0002$	$0,0029 \pm 0,0000$	$0,0029 \pm 0,0000$
	HN	$0,0078 \pm 0,0006$	$0,0019 \pm 0,0002$	$0,0019 \pm 0,0002$
	SC	$87,9 \pm 1,7$	$13,4 \pm 0,2$	$8,8 \pm 0,1$
	MA	$87,9 \pm 1,7$	$13,4 \pm 0,2$	$8,8 \pm 0,1$
	CE	$85,0 \pm 0,6$	$94,5 \pm 0,6$	$95,2 \pm 0,6$
	PC	$98,8 \pm 1,9$	$98,8 \pm 1,9$	$98,8 \pm 1,9$
	AE	$79,06 \pm 2,2$	$8,3 \pm 0,3$	$4,9 \pm 0,1$
	CH	$78,2 \pm 0,7$	$7,7 \pm 0,0$	$4,8 \pm 0,0$
	AN	$0,0116 \pm 0,0002$	$0,0037 \pm 0,0000$	$0,0029 \pm 0,0000$
	HN	$0,0078 \pm 0,0006$	$0,0025 \pm 0,0002$	$0,0019 \pm 0,0002$





Obr. 6: Vliv metody škálování, demonstrováný pomocí GC×GC separace nafty (nahore), moči (uprostřed) a plasmy (dole), na výslednou ortogonalitu, spočítanou pomocí SC (hodnota uvedená v pravém dolním rohu).

### 4.3 Závěr

WOSEL představuje nový způsob škálování 2D separačního prostoru, který do normalizovaného retenčního prostoru převádí celý prostor vymezený analytickou metodou (v obou dimenzích ohraničený mrtvým časem a koncem metody). Na rozdíl od starších metod škálování je díky tomu rozsah separačního prostoru nezávislý na vzorku použitým k optimalizaci.

Při škálování vzorků s velkým pokrytím separačního prostoru poskytuje WOSEL srovnatelné výsledky jako starší metody škálování. Pro vzorky s nižším pokrytím separačního prostoru však dává hodnoty odpovídající vzhledu chromatogramu. Umožňuje optimalizaci analytických systémů i pro analyty, které nejsou ve vzorku momentálně přítomny, přestože analytická metoda pro ně v separaci ponechává prostor.

## 5 Význam dimenzionality vzorku pro výpočet ortogonalita 2D separací

Dimenzionalita vzorku je souhrnem všech nezávislých vlastností vzorku, které mají vliv na jeho chromatografické chování. Většina vzorků jsou vzorky vysokodimenzionální, jejichž dimenzionalita je mnohonásobně vyšší než dimenzionalita použitého chromatografického systému. Píky těchto vzorků vyplňují bez jakéhokoli řádu celý 2D separační prostor. Oproti tomu vzorky nízkodimenzionální mají srovnatelnou dimenzionalitu jako analytický systém. Jejich píky vyplňují jen část chromatogramu. Formují se do skupin podle příslušnosti do chemických tříd. Poloha píku v rámci skupiny je dána pozicí funkční skupiny ve struktuře dané látky [29].

Ortogonalita je parametr, který se používá k optimalizaci chromatografických 2D systémů. Počítá se pomocí celé řady přístupů, pracujících na rozličných principech. Většinou nabývá hodnot od 0 do 100 %. Obecně lze říci, že čím vyšší je ortogonalita (resp. průměrná vzdálenost NN), tím větší přínos má druhá dimenze pro celkovou separaci.

Ortogonalita systému se hodnotí ze dvou různých hledisek, tzv. aspektů. Lokální ortogonalita popisuje rovnoměrnost pokrytí separačního prostoru, zatímco globální ortogonalita rozsah využitého separačního prostoru. V současné době neexistuje deskriptor, který by uměl zachytit obě hlediska ortogonalita. Proto některé deskriptory hodnotí pouze míru nahlučení dat, jiné zase zohledňují, kolik separačního prostoru je pokryto datovými body. Obě hlediska separace jsou však stejně důležitá, proto Schure a Davis v předchozí práci zabývající se porovnáním způsobů výpočtu ortogonalita zkombinovali nejlepší deskriptor globální (CH) a lokální ortogonalita (jednu z metod geometrického pokrytí separačního prostoru, založenou na počítání obsazených binů [30]) tak, že je mezi sebou vynásobili. Tento kombinovaný deskriptor dle závěru jejich práce [27] poskytuje výsledky, které nejlépe odpovídají vzhledu chromatogramů. Ve své práci však u všech přístupů, vyjma CH, použili Min-max škálování, které zkresluje reálný separační prostor. CH tak v jejich práci slouží pouze jako korekce na reálný separační prostor, nezakreslený Min-max škálováním. Druhý parametr, který nebyl v práci Schure a Davise zohledněn, byl případný vliv dimenzionalita vzorku.

Cílem práce je nalezení „univerzálního“ deskriptoru ortogonalita, který by byl jasnou volbou při optimalizaci 2D systému pro analýzu jak vysoko- tak i nízkodimenzionálních vzorků. Každý z deskriptorů, které jsou v této práci porovnávány, měří jinou vlastnost chromatogramu: korelaci mezi retenčními časy první a druhé dimenze [21], poměr počtu obsazených binů k prázdným [20] a jejich kombinaci [24]; poměr plochy obsazené píky k celkové využitelné ploše chromatogramu [23], průměrnou vzdálenost nejbližších sousedů [25], distribuci bodů podél čtyř os protínajících separační prostor [26] a míru shody distribuce

bodů predikované na základě retenčních časů první dimenze se skutečnou distribucí bodů v dimenzi druhé [22]. Je tedy velmi těžké posoudit, který z deskriptorů je nejlepší.

## 5.1 Metody

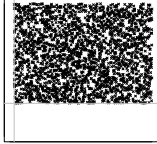
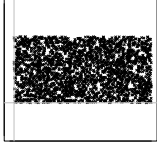

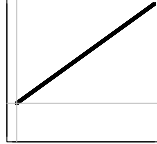
Syntetické chromatogramy jsou v R-software vytvořeny v tak, že 2500 datových bodů různým způsobem vyplňuje separační prostor o rozměrech 89-3000 s × 1-4 s. Algoritmy pro generování syntetických chromatogramů jsou obdobné těm, které byly použity v kapitole 4. K ověření výsledků syntetických chromatogramů je použito reálných GC×GC separací močí, plasmy a nafty. Mrtvé časy jsou určeny pomocí software Column calculator a Chroma-TOF (v. 4.51.6.0, LECO, St. Joseph, Michigan, USA), a to na základě rozměrů použitých kolon, průtoku helia a teplotního programu). Mrtvý čas u syntetických chromatogramů je odvozen na základě reálných separací metabolomických vzorků, připravených ethoximací a trimethylsilylací. Pro přípravu moči se využívá EX (viz kapitola 3), u plasmy deproteinace pomocí methanolu, odfoukání pod dusíkem, ethoximace a trimethylsilylace [31]. Retenční koordináty píků jsou škálovány pomocí WOSEL. Ortogonalita je počítána stejnými přístupy jako v kapitole 4.

## 5.2 Výsledky a diskuse

Správný deskriptor je schopen pokrýt celý rozsah hodnot ortogonalit od 0 do 100 % a umí odlišit průměrné separace od extrémních. Zatímco všechny deskriptory dosáhly ortogonalit blízké nule (Tab. 2, modely C, D), hodnoty 100 % (model A) nedosáhl AE, CE ani HN. Pokles prostoru vyplněného body o 1/3 (model B) odhalily všechny přístupy, až na PC a CE. Lze tedy říci, že AE, CE, HN a PC nereflktují pokrytí separačního prostoru, a proto jsou vyloučeny z dalšího testování. Stejně tak i MA, který je kombinací PC a SC.

Globální ortogonalitu nízkodimenzionálních vzorků nejlépe popisuje AN, který poskytuje nejvyšší hodnotu pro model se 4 shluky umístěnými v rozích separačního prostoru (Obr. 7). CH pro tento model vypočítal ortogonalitu 98,5 %, protože není schopen reflektovat prázdné prostory mezi shluky. Není proto vhodným deskriptorem globální ortogonalit. Nejlepším deskriptorem lokální ortogonalit je SC, protože odráží míru překryvu shluků (Obr. 7). Kombinace AN a SC do finálního deskriptoru pomocí násobení, navrhovaná Schurem a Davisem [27], není nápomocná, pouze zesiluje společný trend jednotlivých deskriptorů. Proto jsou AN a SC zkombinovány prostřednictvím geometrického průměru do ASCA (z angl. „AN a SC average“). ASCA vychází z předpokladu, že rovnoměrné pokrytí a maximální využití separačního prostoru je stejně důležité.

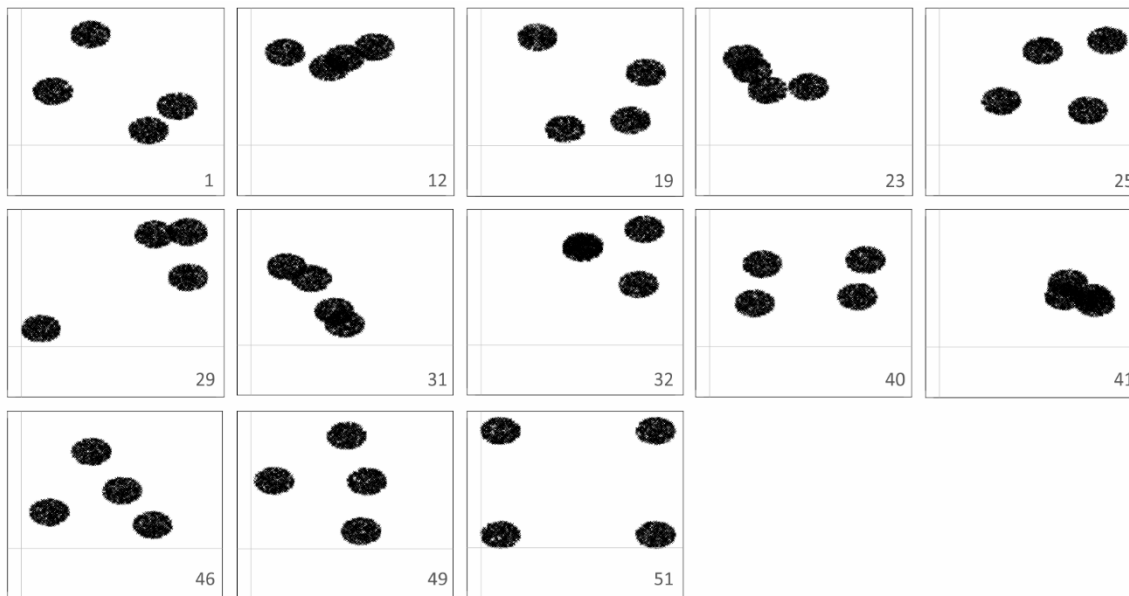
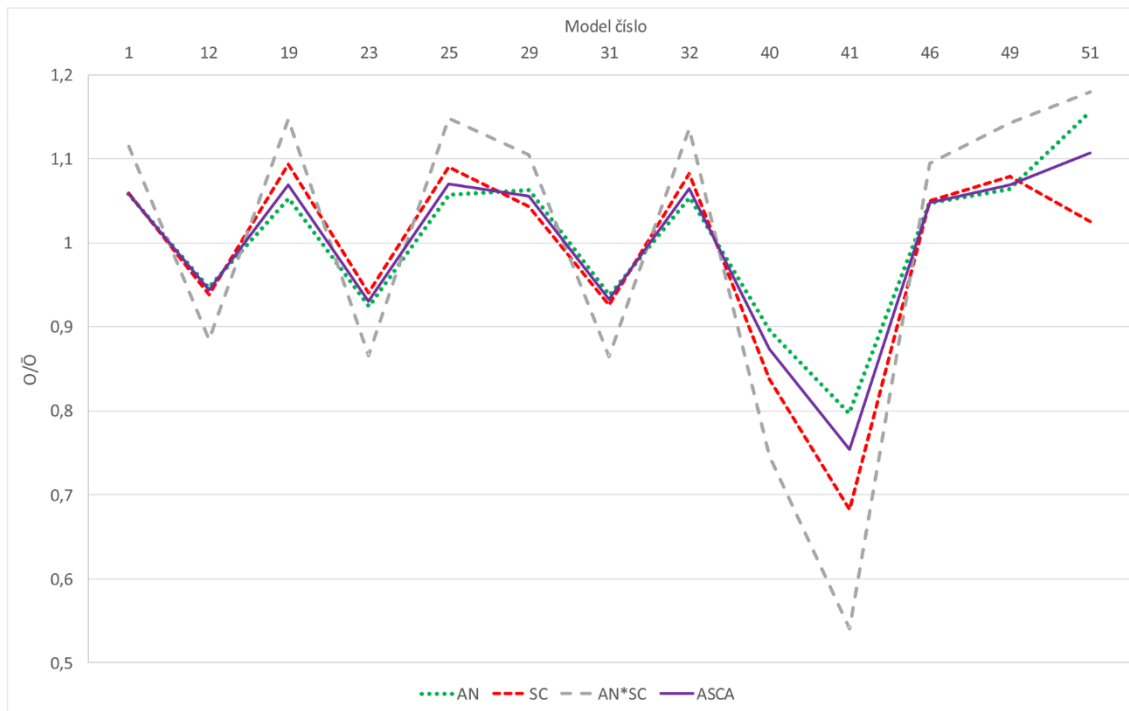
Tab. 2: Výsledky ortogonality [%] ( $\pm 2$  SD pro modely s náhodným rozdělením) poskytnuté jednotlivými přístupy výpočtu pro modely pro určení rozsahu deskriptorů.

Typ	A	B	C	D
Vzhled				
Rozdělení	Náhodné	Náhodné	Pravidelné	Pravidelné
Přístup	PC	98,8 $\pm$ 2,0	98,6 $\pm$ 2,2	96,5
	CE	85,8 $\pm$ 0,6	89,2 $\pm$ 0,8	-
	AN	99,9 $\pm$ 1,2	81,6 $\pm$ 1,1	3,1
	HN	67,3 $\pm$ 3,8	54,6 $\pm$ 3,6	3,1
	CH	99,2 $\pm$ 0,4	66,1 $\pm$ 0,3	0
	AE	96,3 $\pm$ 2,5	70,8 $\pm$ 0,8	0
	MA	100,3 $\pm$ 1,9	82,6 $\pm$ 2,0	3,1
	SC	100,3 $\pm$ 1,9	82,7 $\pm$ 2,0	3,1

U vysokodimenzionálních vzorků je situace zcela jiná. Celý separační prostor je pokryt datovými body, globální ortogonalita proto není informativním parametrem. Optimalizovat lze pouze lokální ortogonalitu, a to nejlépe prostřednictvím SC. Jelikož se však AN chová jako konstanta, lze k optimalizaci použít i ASCA. Ne všechny dimenze (vlastnosti či funkční skupiny) vzorku jsou nezbytné k dosažení cíle analýzy. Ty, které nejsou podstatné, je možné při přípravě vzorku či při analýze dat potlačit. Tak lze i vysokodimenzionální vzorek optimalizovat jako nízkodimenzionální.

### 5.3 Závěr

K optimalizaci nízko- i vysokodimenzionálních vzorků je nevhodnější kombinovaný deskriptor ASCA, který v sobě prostřednictvím geometrického průměru spojuje oba aspekty ortogonality: globální (zastoupenou AN) i lokální (představovanou SC). Dimenzionalita vzorku může být před analýzou snížena s ohledem na zamýšlený cíl.



Obr. 7: Porovnání AN, SC a kombinovaných deskriptorů AN\*SC a ASCA pomocí nízkodimenzionálních syntetických modelů. Nahoře srovnávací graf s hodnotami ortogonality normalizovanými průměrnou hodnotou ortogonality pro jednotlivé přístupy, dole chromatogramy jednotlivých modelů.

## 6 Nalezení nejvíce ortogonální kombinace kolon pro GC×GC analýzu metabolomu moči

Výběr kolon pro GC×GC se řídí snahou separovat co nejvyšší počet komponent vzorku. Proto se kombinují kolony co nejvíce rozdílné (ortogonální), lišící se separačními mechanismy. Tzv. normální kolonové uspořádání je tvořeno nepolární kolonou v první dimenzi, která separuje látky podle disperzních sil (eluční pořadí pak koresponduje s bodem varu látek), a polární (více selektivní) kolonou v dimenzi druhé, jež separuje koelující látky první dimenze podle polarit. Například při analýze nafty první nepolární dimenze umožní separaci alkanů a cykloalkanů, zatímco areny jsou díky  $\pi$ - $\pi$  interakcím zadržovány druhou dimenzí. V tzv. reverzním kolonovém uspořádání je tomu naopak [32].

Ve většině metabolomických aplikací převládá tradiční kolonové uspořádání, tvořené nepolární 100% polymethylsiloxanovou kolonou v první dimenzi a středně polární polymethylsiloxanovou kolonou s 50 % fenylu v dimenzi druhé. Nepolární kolony typu BPX5 se v první dimenzi používají díky dlouholetým zkušenostem z 1D GC, polární kolony druhé dimenze nabízejí jinou selektivitu.

Cílem této části disertační práce je pomocí ortogonálního přístupu (deskriptoru ASCA) otestovat jiné druhy stacionárních fází a nalézt nejlepší kolonové uspořádání pro analýzu metabolomu moči a séra, u nejlepších kolonových uspořádání také určit vliv změny vnitřního průměru kolony druhé dimenze na ortogonalitu systému.

### 6.1 Metody

K stanovení ortogonality kolonových uspořádání je použito směsné moči, získané smícháním moči zdravých dobrovolníků, a plasmy SRM 1950 [33]. Vzorek moči je připraven pomocí EX, uvedené v kapitole 3, vzorek plasmy podle McGaw a kol. [31], prostřednictvím Pegasus 4D GC×GC-TOF MS (LECO) analyzován pomocí různých kolonových uspořádání (Tab. 3, 4). Analýza dat je provedena pomocí ChromaTOF software v. 4.24 (LECO), MS Excel, R-software (v. 3.3.0), and MATLAB (v. R2018a).

### 6.2 Výsledky a diskuse

Většina GC×GC analýz moči připravené pomocí trimethylsilylace je prováděna pomocí normálního kolonového uspořádání, jmenovitě pomocí setu BPX5/BPX50 nebo jejich komerčních analog. O použití jiných kolonových uspořádání k analýze metabolitů moči připravených trimethylsilylací není v literatuře žádná informace. Proto je vzorek moči analyzován pomocí šesti různých kombinací kolon: dvou kombinací kolony nepolární s polární (BPX5/SolGel-WAX, BPX5/SLB-IL59), dvou reverzních setů kolon polárních

s nepolárními (BPX50/BPX5, SolGel-WAX/BPX-5), jednoho uspořádání středně polární kolony s polární kolonou (BPX50/SolGel-WAX) a jedné kombinace polární kolony se středně polární (SLB-IL59/BPX50). Instalace polární kolony v první dimenzi není účelná. U kolony SLB-IL59 (set A) dochází k extrémní retenci analytů (a téměř žádné separaci). Kolona SolGel-WAX (set M) zase extrémně zadržuje rozpouštědla. Kolona SLB-IL59 (set B) není vhodná ani jako kolona druhé dimenze, protože způsobuje extrémní chvostování píků.

Ortogonalita zbývajících kolonových uspořádání spočítaná pomocí ASCA je uvedena v Tab. 3. Všechny sety obsahující v druhé dimenzi kolonu s 0,25  $\mu\text{m}$  stacionární fáze vykazují, ve srovnání se stejným typem stacionární fáze v tenčí vrstvě, vyšší orthogonalitu. Z tohoto jevu lze usuzovat, že silnější vrstva stacionární fáze má výrazný vliv na retenci analytů. Jako nejvíce ortogonální je hodnoceno kolonové uspořádání BPX50/SolGel-WAX (set J,  $O = 68,5 \%$ ).

Tab. 3: Kolonová uspořádání použitá pro analýzu moči a hodnota ortogonalit spočítaná pomocí ASCA

Set	Kolonová uspořádání ( $^1D/{}^2D$ )	Parametry kolony druhé dimenze (délka [m] – průměr [mm] – film [ $\mu\text{m}$ ])	O [%]
A	SLB-IL59/BPX50	1,3 – 0,25 – 0,20	-
B	BPX5/SLB-IL59	2,0 – 0,10 – 0,08	-
C	BPX5/SolGel-WAX	1,2 – 0,10 – 0,10	56,0
H	BPX50/BPX5*	1,2 – 0,10 – 0,10	45,1
I	BPX50/SolGel-WAX	1,6 – 0,10 – 0,10	44,3
J	BPX50/SolGel-WAX	1,2 – 0,25 – 0,25	68,5
K	BPX5/SolGel-WAX	1,2 – 0,25 – 0,25	65,9
L	BPX50/BPX5*	1,3 – 0,25 – 0,25	64,3
M	SolGel-WAX/BPX-5*	1,3 – 0,10 – 0,10	-
N	BPX5/BPX50	1,5 – 0,10 – 0,10	56,9

\* Modulace na koloně druhé dimenze

GC $\times$ GC analýza plasmy je kromě tradičního kolonového uspořádání odpovídajícímu chemismu stacionárních fází BPX5/BPX50 prováděna také pomocí reverzního uspořádání BPX50/BPX5, anebo pomocí RTX200, která obsahuje středně polární trifluoropropylmethylpolysiloxan. Trifluoropropylmethylpolysiloxan interaguje s volnými elektronovými páry a  $\pi$ -elektronovými systémy, proto kolona zvýšeně zadržuje nenasycené sloučeniny, aromáty, aldehydy, ketony, halogenderiváty a nitrosločeniny [34]. Použití iontových kapalin ani polyethylenglykolových kolon pro analýzu trimethylsilylovaných metabolitů moči literaturou doloženo není, proto jim v této práci byla věnována pozornost.

Stejně jako je tomu u analýzy moči, kolony druhé dimenze se silnější vrstvou stacionární fáze vykazují v porovnání se tenčí vrstvou vyšší orthogonalitu (Tab. 4). Jako nejlepší, co se ortogonalit týče, je

hodnocen set BPX50/BPX5. Tento závěr je v souladu s prací Koek a kol. [35]. Polyethylenglykolová kolona ani kolona na bázi iontové kapaliny není pro separaci v normálním uspořádání přínosem. SLB-IL59 u kolony druhé dimenze také u plasm způsobuje chvostování píků. Při použití kolony SolGel-WAX ve druhé dimenzi se na chromatogramu objevuje méně píků s vyšší retencí na této koloně. Je tomu tak nejspíš proto, že plasma je ve srovnání s močí celkově jednodušší materiál (obsahuje méně metabolitů detekovatelných pomocí GC-MS a v nižší koncentraci), a tak k nekompletní derivatizaci dochází zřídka. Analytů obsahujících kyselý vodík, které by mohly interagovat s polyethylenglykolovou stacionární fází, je proto velmi málo. Reverzní uspořádání, kdy by SolGel-WAX a SLB-IL59 byly kolonami první dimenze, nejsou na základě výsledků analýzy moči v případě plasm testovány.

Tab. 4: Kolonová uspořádání použitá pro analýzu plasmy

Set	Kolonová uspořádání ( <sup>1</sup> D/ <sup>2</sup> D)	Parametry kolony druhé dimenze (délka [m] – průměr [mm] – film [μm])	O [%]
A	BPX5/BPX50	1,5 – 0,10 – 0,10	51,0
B	BPX5/SLB-IL59	2,0 – 0,10 – 0,08	53,5
C	BPX50/SolGel-WAX	1,6 – 0,10 – 0,10	44,7
D	BPX50/SolGel-WAX	1,2 – 0,25 – 0,25	56,7
E	BPX5/SolGel-WAX	1,2 – 0,25 – 0,25	58,4
F	BPX50/BPX5*	1,3 – 0,25 – 0,25	62,4
G	BPX5/SolGel-WAX	1,2 – 0,10 – 0,10	40,6

\* Modulace na koloně druhé dimenze

### 6.3 Závěr

Nejvyšší ortogonalitu pro GC×GC analýzu moči poskytuje set BPX50/SolGel-WAX, zatímco pro analýzu plasmy je jím set BPX50/BPX5. V obou případech znamená použití kolony druhé dimenze s větším průměrem a tlustší vrstvou stacionární fáze výrazný nárůst retence analytů a tím i nárůst ortogonality systému.

To, že se nejvíce ortogonální uspořádání kolon pro plasmu a moč liší, je pravděpodobně způsobeno rozdílným charakterem obou biologických materiálů. Protože koncentrace analytů v plasmě je řízena homeostatickými mechanismy, rozdíl v koncentraci analytů není tak velký. Podíl nedostatečně derivatizovaných analytů kvůli nadměrné spotřebě derivatizačního činidla je velmi malý. Proto mají analyty méně kyselých vodíků, které by mohly interagovat s polyethylenglykolovou stacionární fází. Hlavní odlišující vlastností je kromě velikosti molekuly přítomnost π-elektronů.

Oproti tomu moč obsahuje řadu látek, které se do ní jako nepotřebné zplodiny metabolismu z těla vylučují ve vysoké koncentraci. Příkladem je močovina, hydroxyhippuráty aj. Proto je při derivatizaci



metabolitů v moči spotřeba derivatizačního činidla vyšší, stejně jako zastoupení neúplně derivatizovaných látek s kyselými vodíky. SolGel-WAX kolona nainstalovaná v druhé dimenzi je proto pro analýzu moči výhodou. Při současném použití kolony BPX50 jsou první dimenzi jsou analyty separovány na základě Londonových disperzních sil a  $\pi$ - $\pi$  interakcí (tedy zcela odlišně).

## 7 Seznam použité literatury

- [1] N. Vinayavekhin, A. Saghatelian, Untargeted Metabolomics, in: Curr. Protoc. Mol. Biol., John Wiley & Sons, Inc., Hoboken, NJ, USA, 2010: pp. 30.2.1-30.2.24. doi:10.1002/0471142727.mb3001s90.
- [2] W.W. Wells, T. Chin, B. Weber, Quantitative analysis of serum and urine sugars by gas chromatography, Clin. Chim. Acta. 10 (1964) 352–359. doi:10.1016/0009-8981(64)90066-X.
- [3] K. Tanaka, A. West-Dull, D.G. Hine, T.B. Lynn, T. Lowe, Gas-chromatographic method of analysis for urinary organic acids. II. Description of the procedure, and its application to diagnosis of patients with organic acidurias., Clin. Chem. 26 (1980) 1847–53. <http://www.ncbi.nlm.nih.gov/pubmed/7438430>.
- [4] J.D. Shoemaker, W.H. Elliott, Automated screening of urine samples for carbohydrates, organic and amino acids after treatment with urease, J. Chromatogr. B Biomed. Sci. Appl. 562 (1991) 125–138. doi:10.1016/0378-4347(91)80571-S.
- [5] B.-J. Webb-Robertson, Y.-M. Kim, E.M. Zink, K.A. Hallaian, Q. Zhang, R. Madupu, K.M. Waters, T.O. Metz, A Statistical Analysis of the Effects of Urease Pre-treatment on the Measurement of the Urinary Metabolome by Gas Chromatography-Mass Spectrometry., Metabolomics. 10 (2014) 897–908. doi:10.1007/s11306-014-0642-1.
- [6] I. Matsumoto, T. Kuhara, A new chemical diagnostic method for inborn errors of metabolism by mass spectrometry—rapid, practical, and simultaneous urinary metabolites analysis, Mass Spectrom. Rev. 15 (1996) 43–57. doi:10.1002/(SICI)1098-2787(1996)15:1<43::AID-MAS3>3.0.CO;2-B.
- [7] T. Kuhara, Diagnosis of inborn errors of metabolism using filter paper urine, urease treatment, isotope dilution and gas chromatography-mass spectrometry, J Chromatogr B Biomed Sci Appl. 758 (2001) 3–25.
- [8] D.D. van Slyke, G.E. Cullen, Mode of Action of Urease, J. Biol. Chem. 19 (1914) 141–180.
- [9] J.F. Marlier, L.I. Robins, K.A. Tucker, J. Rawlings, M.A. Anderson, W.W. Cleland, A Kinetic and Isotope Effect Investigation of the Urease-Catalyzed Hydrolysis of Hydroxyurea, Biochemistry. 49 (2010)

8213–8219. doi:10.1021/bi100890v.

- [10] R. Karlíková, J. Šíroková, D. Friedecký, E. Faber, M. Hrdá, K. Mičová, I. Fikarová, A. Gardlo, H. Janečková, I. Vrobel, T. Adam, Metabolite Profiling of the Plasma and Leukocytes of Chronic Myeloid Leukemia Patients, *J. Proteome Res.* 15 (2016) 3158–3166. doi:10.1021/acs.jproteome.6b00356.
- [11] E.C.Y. Chan, K.K. Pasikanti, J.K. Nicholson, Global urinary metabolic profiling procedures using gas chromatography–mass spectrometry, *Nat. Protoc.* 6 (2011) 1483–1499. doi:10.1038/nprot.2011.375.
- [12] T. Kind, V. Tolstikov, O. Fiehn, R.H. Weiss, A comprehensive urinary metabolomic approach for identifying kidney cancer, *Anal. Biochem.* 363 (2007) 185–195. doi:10.1016/j.ab.2007.01.028.
- [13] M. Jořenek, Studium biologicky aktivních látek imobilizovaných na magnetických nosičích a jejich použití v biotechnologických procesech, Univerzita Palackého v Olomouci, 2016.
- [14] A. Gardlo, AlzbetaG/Metabol: First version, (2019). doi:10.5281/zenodo.3235775.
- [15] G.F. Hoffmann, F.K. Trefz, P.G. Barth, H.J. Böhles, B. Biggemann, H.J. Bremer, E. Christensen, M. Frosch, F. Hanefeld, D.H. Hunneman, Glutaryl-coenzyme A dehydrogenase deficiency: a distinct encephalopathy., *Pediatrics.* 88 (1991) 1194–203. <http://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed2&NEWS=N&AN=1992029708>.
- [16] G.L. Hedlund, N. Longo, M. Pasquali, Glutaric acidemia type 1., *Am. J. Med. Genet. C. Semin. Med. Genet.* 142C (2006) 86–94. doi:10.1002/ajmg.c.30088.
- [17] W. Steuer, I. Grant, F. Erni, Comparison of high-performance liquid chromatography, supercritical fluid chromatography and capillary zone electrophoresis in drug analysis., *J. Chromatogr.* 507 (1990) 125–40. doi:10.1016/S0021-9673(01)84188-1.
- [18] P.J. Slonecker, X. Li, T.H. Ridgway, J.G. Dorsey, Informational orthogonality of two-dimensional chromatographic separations., *Anal. Chem.* 68 (1996) 682–9. doi:10.1021/ac950852v.
- [19] J.M. Davis, D.R. Stoll, P.W. Carr, Effect of first-dimension undersampling on effective peak capacity in comprehensive two-dimensional separations., *Anal. Chem.* 80 (2008) 461–73. doi:10.1021/ac071504j.
- [20] M. Gilar, P. Olivova, A.E. Daly, J.C. Gebler, Orthogonality of separation in two-dimensional liquid chromatography., *Anal. Chem.* 77 (2005) 6426–34. doi:10.1021/ac050923i.
- [21] E. Van Gysegheem, I. Crosiers, S. Gourvénec, D. Massart, Y. Vander Heyden, Determining orthogonal and similar chromatographic systems from the injection of mixtures in liquid chromatography–

- diode array detection and the interpretation of correlation coefficients color maps, *J. Chromatogr. A.* 1026 (2004) 117–128. doi:10.1016/j.chroma.2003.10.116.
- [22] M.R. Pourhaghighi, M. Karzand, H.H. Girault, Orthogonality of Two-Dimensional Separations Based on Conditional Entropy, *Anal. Chem.* 83 (2011) 7676–7681. doi:10.1021/ac2017772.
- [23] G. Semard, V. Peulon-Agasse, A. Bruchet, J.-P.P. Bouillon, P. Cardinaël, Convex hull: A new method to determine the separation space used and to optimize operating conditions for comprehensive two-dimensional gas chromatography, *J. Chromatogr. A.* 1217 (2010) 5449–54. doi:10.1016/j.chroma.2010.06.048.
- [24] Z.-D. Zeng, H.M. Hugel, P.J. Marriott, A modeling approach for orthogonality of comprehensive two-dimensional separations., *Anal. Chem.* 85 (2013) 6356–63. doi:10.1021/ac400736v.
- [25] W. Nowik, S. Héron, M. Bonose, M. Nowik, A. Tchapla, Assessment of two-dimensional separative systems using nearest-neighbor distances approach. Part 1: orthogonality aspects., *Anal. Chem.* 85 (2013) 9449–58. doi:10.1021/ac4012705.
- [26] M. Camenzuli, P.J. Schoenmakers, A new measure of orthogonality for multi-dimensional chromatography., *Anal. Chim. Acta.* 838 (2014) 93–101. doi:10.1016/j.aca.2014.05.048.
- [27] M.R. Schure, J.M. Davis, Orthogonal separations: Comparison of orthogonality metrics by statistical analysis, *J. Chromatogr. A.* 1414 (2015) 60–76. doi:10.1016/j.chroma.2015.08.029.
- [28] W. Hoeffding, Scale—Invariant Correlation Theory, in: 1994: pp. 57–107. doi:10.1007/978-1-4612-0865-5\_4.
- [29] J.C. Giddings, Sample dimensionality: a predictor of order-disorder in component peak distribution in multidimensional separation., *J. Chromatogr. A.* 703 (1995) 3–15. doi:10.1016/0021-9673(95)00249-M.
- [30] M.R. Schure, The dimensionality of chromatographic separations, *J. Chromatogr. A.* 1218 (2011) 293–302. doi:10.1016/j.chroma.2010.11.016.
- [31] E.A. McGaw, K.W. Phinney, M.S. Lowenthal, Comparison of orthogonal liquid and gas chromatography–mass spectrometry platforms for the determination of amino acid concentrations in human plasma, *J. Chromatogr. A.* 1217 (2010) 5822–5831. doi:10.1016/j.chroma.2010.07.025.
- [32] R.C.Y. Ong, P.J. Marriott, A Review of Basic Concepts in Comprehensive Two-Dimensional Gas Chromatography, *J. Chromatogr. Sci.* 40 (2002) 276–291. doi:10.1093/chromsci/40.5.276.
- [33] K.W. Phinney, G. Ballihaut, M. Bedner, B.S. Benford, J.E. Camara, S.J. Christopher, W.C. Davis, N.G. Dodder, G. Eppe, B.E. Lang, S.E. Long, M.S. Lowenthal, E.A. McGaw, K.E. Murphy, B.C. Nelson, J.L. Prendergast, J.L. Reiner, C.A. Rimmer, L.C. Sander, M.M. Schantz, K.E. Sharpless, L.T. Sniegoski, S.S.-

C. Tai, J.B. Thomas, T.W. Vetter, M.J. Welch, S.A. Wise, L.J. Wood, W.F. Guthrie, C.R. Hagwood, S.D. Leigh, J.H. Yen, N.-F. Zhang, M. Chaudhary-Webb, H. Chen, Z. Fazili, D.J. LaVoie, L.F. McCoy, S.S. Momin, N. Paladugula, E.C. Pendergrast, C.M. Pfeiffer, C.D. Powers, D. Rabinowitz, M.E. Rybak, R.L. Schleicher, B.M.H. Toombs, M. Xu, M. Zhang, A.L. Castle, Development of a Standard Reference Material for metabolomics research., *Anal. Chem.* 85 (2013) 11732–8. doi:10.1021/ac402689t.

[34] W. Engewald, K. Dettmer-Wilde, H. Rotzsche, Columns and Stationary Phases, in: *Pract. Gas Chromatogr.*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2014: pp. 59–116. doi:10.1007/978-3-642-54640-2\_3.

[35] M.M. Koek, B. Muilwijk, L.L.P. van Stee, T. Hankemeier, Higher mass loadability in comprehensive two-dimensional gas chromatography–mass spectrometry for improved analytical performance in metabolomics analysis, *J. Chromatogr. A.* 1186 (2008) 420–429. doi:10.1016/j.chroma.2007.11.107.

# ŽIVOTOPIS

## Osobní údaje

Jméno: Jaroslava Jáčová  
Adresa: Praskova 446/5, 779 00 Olomouc  
Email: jaroslava.jacova@gmail.com

## Vzdělání

2013 – současnost Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Katedra analytické chemie, doktorské studium oboru Analytická chemie

2016 Institut postgraduálního vzdělávání ve zdravotnictví – IPVZ, Akreditovaný kvalifikační kurz Odborné zdravotnické laboratorní metody

2012 Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Katedra analytické chemie, státní rigorózní zkouška  
Rigorózní práce: „Stanovení lanthanu v křemičitém skle“

2004 – 2006 Ostravská univerzita, Pedagogická fakulta, Centrum dalšího vzdělávání, Doplňující pedagogické studium pro učitele odborných předmětů.  
Závěrečná práce: „Hodnotová orientace žáků 2. ročníku SUPŠ sklářské“

1999 – 2001 Univerzita Pardubice, Fakulta chemicko-technologická, Katedra biologických a biochemických věd, magisterské studium oboru Analýza biologických materiálů  
Diplomová práce: „Adaptace lidských biologických hodin k přechodu na letní čas“

1996 – 1999 Univerzita Pardubice, Fakulta chemicko-technologická, Katedra biologických a biochemických věd, bakalářské studium oboru Klinická biologie a chemie  
Bakalářská práce: „Melatonin – funkce v lidském organismu, metody stanovení“

## Stáže

1. 6. – 31. 8. 2017 Slovenská akadémia vied, Neuroimunologický ústav

## Projekty

2019	Molekulární podstata a diagnostické markery vybraných hematologických a systémových chorob X., UPOL, IGA_LF_2019_006
2018 – nyní	Charakterizace lidského lipidomu a metabolomu pro personalizovanou zdravotní péči a hledání biomarkerů: studie rakoviny ledvin, GAČR, projekt. č. GA18-12204S
2018	Molekulární podstata a diagnostické markery vybraných hematologických a systémových chorob IX., UPOL, IGA_LF_2018_010
2015 – 2017	Statistika v metabolomice pro výzkum biomarkerů v medicíně, GAČR, projekt č. GF15-34613L
2014 – 2019	Podpora udržitelnosti Ústavu molekulární a translační medicíny, MŠMT, projekt č. NPU I LO1304
2012 – 2013	Soubor materiálů k badatelským aktivitám žáků ZŠ a SŠ v přírodních vědách, MŠMT, projekt č. CZ.1.07/1.1.00/26.0032
2008 – 2010	Průmyslový výzkum agregace skelných materiálů a jejich povrchové zpevnění pro využití v interiéru, MPO ČR, projekt č. FI-IM5/050
2006 – 2009	Impuls MSK, MPO ČR, projekt č. FI-IM3/177
1999 - 2001	GLASS CENTRUM Valašské Meziříčí, MPO ČR, projekt č. FB-C2/96

## Pedagogická činnost:

2018 – nyní	vedoucí diplomové práce: „Analýza mastných kyselin s krátkým řetězcem v plasmě pacientek s gestačním diabetem“, Bc. Eliška Ivanovová, Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Laboratoř růstových regulátorů
2017 – 2018	vedoucí diplomové práce: „Imobilizovaná ureasa pro metabolomiku lidské moči založenou na GC/MS“, Bc. Nikola Čipáková, Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Katedra biochemie
2017 – 2018	vedoucí bakalářské práce: „Vliv pH na aktivitu ureasy pro GC/MS metabolomiku lidské moči“, Eliška Ivanovová, Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Laboratoř růstových regulátorů
2015 – 2016	vedoucí bakalářské práce: „Analýza sacharidů a polyolů v moči pomocí GC/MS pro diagnostiku metabolických poruch“, Nikola Zdráhalová, Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Katedra fyzikální chemie
2013 – 2014	vedoucí diplomové práce: „Analýza metabolomu lidské moči pomocí GCxGC-TOF“, Bc. Marie Zdráhalové, Univerzita Palackého v Olomouci, Přírodovědecká fakulta, Katedra biochemie

2003 – 2013 učitelka odborných předmětů, Střední umělecko-průmyslová škola sklářská Valašské Meziříčí

### Ocenění

2018 cena děkana LF UP za nejlepší studentskou vědeckou práci za rok 2017  
2017 Best Poster Prize at MetaboMeeting 2017, Birmingham, UK  
2001 cena Nadačního fondu Miroslava Jurečka za vynikající úroveň diplomové práce  
Chemicko-technologické fakulty Univerzity Pardubice

### Publikace v časopisech s impaktním faktorem

2019 J. Jáčová, M. Jořenek, K. Pospíšková, L. Najdekr, L. Zajoncová, D. Friedecký, T. Adam, Urease-immobilized magnetic microparticles in urine sample preparation for metabolomic analysis by gas chromatography-mass spectrometry, J. Chromatogr. A., v recenzním řízení  
2019 J. Jáčová, A. Gardlo, J.-M.D. Dimandja, T. Adam, D. Friedecký, Impact of sample dimensionality on orthogonality metrics in comprehensive two-dimensional separations, Anal. Chim. Acta. (2019). doi:10.1016/j.aca.2019.03.018. IF 5,123  
2017 J. Jáčová, A. Gardlo, D. Friedecký, T. Adam, J.-M.D. Dimandja, Sample-independent approach to normalize two-dimensional data for orthogonality evaluation using whole separation space scaling, J. Chromatogr. A. 1511 (2017) 1–8. doi:10.1016/j.chroma.2017.06.076. IF 3,716

### Ostatní publikace

2015 L. Bartáková, P. Daniš, J. Jáčová, Lukáš Müller, Náměty k mimoškolní činnosti - chemie, Univerzita Palackého v Olomouci 2016, ISBN 978-80-244-4745-2  
2015 L. Bartáková, P. Daniš, J. Jáčová, L. Müller: Laboratorní a terénní cvičení - chemie, Univerzita Palackého v Olomouci 2015, ISBN 978-80-244-4590-8  
2015 L. Bartáková, P. Daniš, J. Jáčová, L. Müller: Čítanka k přírodním vědám - chemie, Univerzita Palackého v Olomouci 2015, ISBN 978-80-244-4509-0  
2011 J. Jáčová, J. Hřebíčková: Laboratorní metody. Univerzita Tomáše Bati ve Zlíně [online], 2011. [cit. 2019-06-05]. Dostupné z WWW: <<http://ads.fmk.utb.cz/Contexts/atelier/Documents/Laboratorn%C3%AD%20metody.pdf>>

## Konference

- 2019 HPLC 2019, Miláno, Itálie, 16. – 18. 6. 2019, přednáška: “Diagnosing of organic acidurias and beta-oxidation defects by LC-MS/MS”
- 2019 34. pracovní dny dědičné metabolické poruchy, Olomouc, Česká republika, 29. 5. – 31. 5. 2019, přednáška: „Diagnostika organických acidurií a poruch  $\beta$ -oxidace mastných kyselin pomocí LC-MS/MS
- 2019 Česká konference hmotnostní spektrometrie 2019, 27.3. – 29. 3. 2019, Olomouc, Česká republika, přednáška: „Diagnosing of organic acidurias and beta-oxidation defects by LC-MS/MS”
- 2018 MetaboMeeting 2018, Nottingham, UK, 17. – 19. 12. 2018, poster: „Profiles of organic acids by liquid chromatography-mass spectrometry”
- 2018 MOVISS, Vorau, Rakousko, 9. – 12. 9. 2018, pasivní účast
- 2018 15th GCXGC Symposium, Riva del Garda, Itálie, 13. 5 – 18. 5. 2018, plenární přednáška „Bottlenecks of orthogonality evaluation”
- 2018 42nd International Symposium on Capillary Chromatography, Riva del Garda, Itálie, 13. 5 – 18. 5. 2018, poster “ „Urease immobilized on magnetic microparticles for GC/MS-based urinary metabolomics”
- 2107 Metabomeeting 2017, Birmingham, UK, 10. – 13. 12. 2017, poster „Urease immobilized on magnetic microparticles for GC/MS-based urinary metabolomics”
- 2017 MOVISS 2017, Vorau, Rakousko, 20. - 22. 9. 2017, pasivní účast
- 2017 41st International Symposium on Capillary Chromatography & 14th GCxGC Symposium, Texas, USA, 14. - 19. 5. 2017, přednáška „The impact of normalization on orthogonality evaluation”
- 2017 6. konference České společnosti pro hmotnostní spektrometrii, Olomouc, Česká republika, 29. - 31. 3. 2017, přednáška „Dimenzionalita vzorku v optimalizaci 2D separací”
- 2016 Annual Symposium of the Society for the Study of Inborn Errors of Metabolism - SSIEM 2016, Řím, Itálie, 6. – 9. 9. 2016, poster: Urease immobilized on magnetic microparticles for urine GC/MS analysis”
- 2016 Advances in chromatography and electrophoresis & Chiranal 2016, Olomouc, Česká republika, 6. 6. – 9. 6. 2016, přednáška „Evaluation of orthogonality assessment tools in GCxGC”.
- 2016 ISCC - 40th ISCC and 13th GCxGC Symposium, Riva del Garda, Itálie, 29. 5. - 3. 6. 2016, vyžádaná přednáška „Evaluation of orthogonality assessment tools in GCxGC”.



- 2016 31. Pracovní dny dědičné metabolické poruchy, Bratislava, Slovenská republika, 25.5. – 27. 5. 2016, přednáška „Ureáza imobilizovaná na magnetických mikročásticích pro metabolomiku lidské moči“.
- 2016 5. konference České společnosti pro hmotnostní spektrometrii, České Budějovice, Česká republika, 13.-15. 4. 2016, přednáška „Ureáza imobilizovaná na magnetických mikročásticích pro metabolomiku lidské moči“.
- 2015 30. Pracovní dny Dědičné metabolické poruchy, Dolní Morava, Česká republika, 20. – 22. 5. 2015, poster „Stanovení galaktitolu v moči pomocí GC/MS“
- 2014 XXXIII. Regionální pracovní dny klinické biochemie, Karlova Studánka, česká republika, 3. – 5. 12. 2014, přednáška „Kontrola kvality v diagnostice metabolických chorob“
- 2014 38th International Symposium on Capillary Chromatography (ISCC) and 11th GCxGC Symposium, Riva del Garda, Itálie, 18. – 23. 5. 2014, poster „Evaluation of different column combinations in GCxGC-TOF MS analysis of organic acids in human urine“