

## Novel spectroscopic biomarkers are applicable in non-invasive early detection and staging classification of colorectal cancer

M. MISKOVICOVA<sup>1</sup>, V. FRYBA<sup>2</sup>, L. PETRUZELKA<sup>1</sup>, V. SETNICKA<sup>3</sup>, A. SYNITSYA<sup>3</sup>, M. TATARKOVIC<sup>3</sup>, J. ULRYCH<sup>3</sup>, M. VOCKA<sup>1</sup>

<sup>1</sup>Department of Oncology, First Faculty of Medicine, Charles University and General University Hospital, Prague, Czech Republic; <sup>2</sup>1st Surgical Department, First Faculty of Medicine, Charles University and General University Hospital, Prague, Czech Republic; <sup>3</sup>Department of Analytical Chemistry, University of Chemistry and Technology Prague, Prague, Czech Republic

\*Correspondence: [michaela.miskovicova@vfn.cz](mailto:michaela.miskovicova@vfn.cz)

Received May 6, 2020 / Accepted May 25, 2020

Colorectal cancer (CRC) remains a significant threat to human health because of the lack of awareness of physical examination or the limitations of an early diagnostic level. Despite the improving standard of modern medicine, mortality from CRC is still remarkably high and the prognosis remains poor in many cases because of disease detection at advanced clinical stages. Raman spectroscopy yields precise information, not only regarding the secondary structure of proteins but also regarding the discrimination between normal and malignant tissues. We investigated whether this method can be used for the diagnosis of CRC including initial stages. To acquire more detailed structural information, we tested a novel diagnostic approach based on a suitable combination of conventional methods of molecular spectroscopy (Raman and Fourier transform infrared) with advanced, highly structure-sensitive chiroptical techniques as electronic circular dichroism (ECD) and Raman optical activity (ROA) to monitor the CRC pathogenesis relating compositional, structural and conformational changes in blood biomolecules, some of which may be caused by pathological processes occurring during cancer growth, also at the beginning of the disease. Sixty-three blood plasma samples were analyzed using the combination of ECD and ROA supplemented by Raman and Fourier transform infrared (FT-IR) spectroscopies. The obtained spectra were evaluated together by linear discriminant analysis. The accuracy of sample discrimination reached 100% and the subsequent leave-one-out cross-validation resulted in 90% sensitivity and 75% specificity. There were also found the differences between the patients according to the clinical stage. The achieved results suggest a panel of promising biomarkers and indicate that chiroptical methods combined with conventional spectroscopies might be a new minimally invasive powerful tool for producing high-quality data, obtaining an accurate diagnosis of colorectal cancer through a peripheral blood sample, which is also able to determine the extent of this pathology. Further work needs to be carried out for these techniques to be implemented in the clinical setting.

*Key words: colorectal cancer, diagnosis, staging, spectroscopy, Raman optical activity, electronic circular dichroism*

Colorectal cancer (CRC) is the third deadliest and the fourth most diagnosed cancer in the world, with 1.8 million new cases and almost 861 000 deaths in 2018 [1]. Incidence has been steadily rising worldwide, especially in developing nations. Cancer arises, when certain cells of the epithelium of the large intestine develop a series of genetic or epigenetic mutations that confer on them a selective advantage [2]. With abnormally heightened replication and survival, these hyperproliferative cells give rise to a benign adenoma, which may then evolve into carcinoma and metastasize over decades [3].

Although it progresses slowly over a long period of time, it is often detected at advanced stages when the prognosis is already poor. CRC often develops without obvious early symptoms, and a substantial proportion of the at-risk

population does not take advantage of screening offers. Colonoscopy, today's gold standard for the early detection and removal of precancerous lesions, is sometimes refused by the patients due to its invasiveness and discomfort, it requires oral bowel cleansing preparation and it might cause colon perforation [4]. Established non-invasive tests, such as fecal occult blood tests, have high specificity but limited sensitivity, especially with respect to the detection of precursors of CRC, such as adenomas. Biopsy or tumor histopathology after resection is used to screen the precancerous and cancerous lesions of colorectum as a gold standard technique with gross limitations [5–7], which is destructive, time-consuming and depends on the visual observation of pathologists, although it is cost-effective, well-targeted, and high quality [8]. There-

fore, there is a need for the discovery of novel minimally invasive methods and biomarkers that are objective, rapid, and highly accurate, which can identify colorectal cancer and its precursors in easily accessible biospecimens in the early and curable stage [9].

New diagnostic methods based on molecular spectroscopy analysis of cancer tissues or biofluids have been investigated. Vibrational spectroscopic techniques such as Raman spectroscopy and infrared spectroscopy (IR) have great potential for the analysis of biological specimens. Raman spectroscopy is a powerful label-free modality that measures scattered light and can provide valuable structural information about materials and biological samples [10]. Its ability to detect variance related to DNA/RNA, proteins, and lipids have made it an excellent tool for quantifying changes on the cellular level, as well as differentiating between the various cell and tissue types [11]. The collection of Raman spectra from biological samples produces a “fingerprint” representing the molecular vibrations specific to chemical bonds, thus yielding information of a sample’s biochemical composition [12]. Therefore, the Raman spectroscopy has immense potential for medicinal applications and by this method, we can detect specific molecular anomalies in biological tissues, body fluids, and even single cells. Due to this advantage, Raman spectroscopy is especially interesting for diagnostic purposes. This method has already demonstrated its potential capabilities in the identification and objective classification of biochemical changes related to carcinogenesis. There have been few studies of vibrational spectroscopic techniques for the advancement of colorectal cancer diagnostics based on the analysis of blood serum [13, 14].

Because biomolecules in human biofluids are already routinely used to determine the biomarkers of several diseases [15], blood plasma seems to be useful also for the screening of CRC. Blood plasma is easy to obtain, with minimal invasiveness and simple sample preparation. Combining these features with relatively cheap, fast, and reagent-free spectroscopic methods may enable the identification of disease-specific molecular signatures based on the spectral pattern recognition [16, 17]. Changes in spectral patterns are usually caused by variations in molecular structure. In the case of plasmatic biomolecules, these structural alterations are triggered by several pathological processes that occur during the pathogenesis of disease [18].

Unfortunately, few of the currently used diagnostic tools can provide information about the structure of plasmatic biomolecules. Similarly, the use of conventional spectroscopic methods, such as Raman or infrared spectroscopy, has many limitations. The high number of fluorescing entities and low concentrations of molecules of interest represent major drawbacks in the spectroscopic analysis of biofluids [16, 17]. Nevertheless, many bodily biomolecules and potential biomarkers are chiral. Thus, their structure and behavior are suited to be monitored by chiroptical spectroscopy, which is inherently sensitive to the 3D arrangement of chiral

molecules and, thus, better able to identify any subtle conformational changes [19, 20].

Spectroscopic chiroptical methods are widely used in structural and conformational analysis of biomolecules including proteins, polysaccharides, nucleic acids, and even intact viruses. Involving these methods into the investigation of complex biological samples like biofluids would provide new insight into the molecular diagnosis of protein misfolding diseases. Chiroptical methods are among the few methods inherently sensitive to the 3D structures of chiral biomolecules. We suppose that they have a potential for the detection of slight conformational and stereochemical changes in the secondary and tertiary structure of biomolecules caused by pathological processes occurring during several diseases, such as cancer [21]. The process of carcinogenesis leads to detectable alterations and deranging of native biological molecules and supramolecular complexes. These chemical and disorder changes could be recognized by modern spectroscopic and chiroptical methods.

In this study, we combine the structure-sensitive methods of chiroptical spectroscopy (ROA and ECD) complemented by conventional Raman and FT-IR spectroscopies, to compare the spectral patterns of plasma peptides/proteins and other blood biomolecules like glycoproteins, lipids, carbohydrates, and carotenoids of patients suffering from CRC with those of subjects in a control group. We believe that this approach allows us to identify alterations that may be synonymous with CRC and suggest a panel of promising biomarkers for reliable detection of CRC as well as early-stage detection and detection of colorectal cancer recurrence.

## Patients and methods

**Subjects.** For this study, 30 patients with CRC and 33 healthy control subjects were recruited from the Department of Oncology of the General University Hospital and the First Faculty of Medicine, Charles University in Prague, Czech Republic. The average age of CRC patients and controls was 65 (median 66) and 52 (median 55) years, respectively.

The patients with CRC were included in the study after the examination according to diagnostic standards using colonoscopy with confirmed histological verification of colorectal carcinoma (by a histopathologist) and the image diagnostic method Computed tomography (CT) scan of the whole body, which is necessary for staging, the determination of the extent of cancer, which reveals cases of pathological changes including tumor spread or metastases. Patients with rectal carcinoma underwent magnetic resonance imaging (MRI) of the pelvis, which is a basic imaging diagnostic method for the locoregional staging of rectal tumors. If the patients have undergone the surgery at the 1<sup>st</sup> Surgical Department of the General University Hospital and the First Faculty of Medicine, Charles University in Prague, the blood was collected just before this treatment. Then the pathological staging was completed according to the definitive

histological examination of the removed tissue – the tumor and the lymph nodes. Patients' characteristics including the conventional histopathological examination of the tumor and clinical or pathological staging are summarized in Table 1 and 2.

The 30 patients who provided samples for the study were 20 males and 10 females. Locating in the caecum (n=3; 10%), ascending colon (n=4; 13%), transverse colon (n=6; 20%), descending colon (n=1; 3%), sigmoid colon (n=7; 23%), rectosigmoid (n=8; 27%) or rectum (n=1; 3%), all the cancers were classified histologically as adenocarcinoma with various grade of differentiation. Two patients (7%) had well-differentiated adenocarcinoma (grade 1), twenty-one patients (70%) had moderately differentiated adenocarcinoma (grade 2) and adenocarcinoma of seven patients (23%) was poorly differentiated (grade 3). Tumor stages according to the UICC (International Union Against Cancer) were as follows: stage I (n=2; 7%), stage II (n=9; 30%), stage III (n=6; 20%), and stage IV (n=13; 43%). The patients were also subdivided into 2 groups according to the stage and the results of these 2 groups were

compared to each other and to the healthy controls. The 1<sup>st</sup> group included patients of stage I+II (n=11), which represents early stages with no presence of regional lymphadenopathy or distant metastases. The 2<sup>nd</sup> group comprised patients with the tumor spread to regional lymph nodes or to the other parts of the body (stage III+IV; n=19).

A group of healthy controls represents individuals, who underwent preventive screening colonoscopy with a negative result, moreover, the healthy subjects were included in the study after a thorough evaluation of their history and clinical examination with no pathological finding. The age of the control group was comparable to the patient group.

Also, from the group of patients, we selected 5 patients who have undergone curative therapy (surgery +/- adjuvant chemotherapy) and after this treatment, they have been 2 years in remission. That means, that they have no signs of cancer relapse according to clinical, laboratory, and imaging methods. We called them "healthy patients" and their spectra were after 2 years repeatedly evaluated and compared with their first results and with the control group.

**Table 1. Patients and tumor characteristics (n=30).**

No	Age (years)	Sex	Tumor location	Morphological type	T	N	M	Stage
1	63	male	sigmoid	adenocarcinoma grade 2	pT2	pN1a	M0	IIIA
2	77	male	rectosigmoid	adenocarcinoma grade 2	pT3	pN0	M0	IIA
3	72	male	sigmoid	adenocarcinoma grade 2	pT4b	pN2a	M0	IIIC
4	60	male	rectosigmoid	adenocarcinoma grade 2	pT4a	pN1c	M1	IV
5	63	male	rectosigmoid	adenocarcinoma grade 2	pT3	pN0	M0	IIA
6	37	male	rectosigmoid	adenocarcinoma grade 3	pT3	pN0	M0	IIA
7	47	male	sigmoid	adenocarcinoma grade 1	pT3	pN0	M0	IIA
8	80	female	sigmoid	adenocarcinoma grade 2	pT3	pN0	M0	IIA
9	64	female	ascending	adenocarcinoma grade 2	pT3	pN0	M0	IIA
10	62	male	transverse	adenocarcinoma grade 3	pT3	pN0	M0	IIA
11	66	female	sigmoid	adenocarcinoma grade 2	pT4b	pN2a	M1	IV
12	52	female	caecum	adenocarcinoma grade 2	pT3	pN1b	M0	IIIB
13	67	male	caecum	adenocarcinoma grade 3	pT3	pN1a	M0	IIIB
14	55	male	sigmoid	adenocarcinoma grade 2	pT1	pN0	M0	I
15	77	female	caecum	adenocarcinoma grade 2	pT2	pN0	M0	I
16	86	female	ascending	adenocarcinoma grade 3	pT3	pN0	M0	IIA
17	66	male	ascending	adenocarcinoma grade 2	pT3	pN1b	M0	IIIB
18	75	male	transverse	adenocarcinoma grade 2	pT3	pN0	M0	IIA
19	49	female	transverse	adenocarcinoma grade 2	pT3	pN0	M0	IIA
20	46	male	transverse	adenocarcinoma grade 3	T4	N1	M1	IV
21	67	female	rectosigmoid	adenocarcinoma grade 2	T3	N1	M1	IV
22	76	female	transverse	adenocarcinoma grade 3	pT4a	pN1a	M1	IV
23	62	female	sigmoid	adenocarcinoma grade 2	pT3	pN1a	M1	IV
24	71	male	rectosigmoid	adenocarcinoma grade 2	T4	N1	M1	IV
25	74	male	ascending	adenocarcinoma grade 2	pT4	pN2a	M1	IV
26	59	male	rectosigmoid	adenocarcinoma grade 1	pT2	pN1a	M1	IV
27	70	male	rectosigmoid	adenocarcinoma grade 2	T4a	N0	M1	IV
28	60	male	transverse	adenocarcinoma grade 3	pT4a	pN1b	M1	IV
29	59	male	rectum	adenocarcinoma grade 2	pT2	pN0	M1	IV
30	75	male	descending	adenocarcinoma grade 2	T3	N1	M1	IV

Blood samples were collected from both studied groups and subjected to a routine biochemical test along with the assessment of human serum albumin, total protein, fasting plasma glucose, and C-Reactive Protein. Cancer biomarker levels (Ca 19-9 and CEA) were measured for both groups (Table 3). The levels of all parameters were much more pathological in patients with CRC than in the control group.

**Ethics statement.** The study was approved by the Ethics Committee of the First Faculty of Medicine and the General University Hospital, Prague. Written informed consent was secured from all subjects.

**Blood plasma.** Whole venous blood (9 ml) was collected by venipuncture from the fasting participants, on an empty stomach and without the use of cigarettes and coffee, and before the administration of a morning medication. The blood samples were obtained at the Department of Oncology, First Faculty of Medicine, Charles University and the General University Hospital in Prague using the sterile anticoagulant-treated blood collection tubes (with K<sub>3</sub>EDTA-tripotassium salt of ethylene diamine tetraacetic acid, BD Vacutainer Systems, Plymouth, UK). Plasma fractions were gained by centrifugation at 1500×g in the temperature 25 °C for 10 min, immediately frozen and stored at -75 °C. Before each analysis at the Department of Analytical Chemistry, University of Chemistry and Technology, Prague, the frozen plasma samples were thawed at room temperature and filtered through a polyvinylidene difluoride (PVDF) membrane filter with 0.45 μm sized pores (Grace, Chicago, IL, USA) at 13,000×g and 15 °C for 10 minutes to remove any residual particulates.

**Infrared spectroscopy.** A Nicolet 6700 (Thermo Scientific, USA) infrared spectrometer with Fourier transformation (FT-IR) was used to measure the sample spectra in the spectral range of 4000–600 cm<sup>-1</sup>. Because of the strong absorption of water, attenuated total reflection (ATR) technique with a ZnSe crystal was used. The filtered samples

(10 μl) were analyzed without any additional pre-treatment. Five hundred and twelve scans were collected to create each individual spectrum with a resolution of 4 cm<sup>-1</sup>. Water and water vapor spectra measured under identical conditions were subtracted from all sample spectra. Linear baseline correction was performed in the OMNIC 32 program, ver. 8.2 (Thermo Scientific, USA).

**Raman optical activity and Raman spectroscopy.** The Raman optical activity and Raman spectra were measured simultaneously using the ChiralRaman-2X<sup>TM</sup> spectrometer (BioTools, Inc., USA) equipped with Laser Quantum MPC6000 Opus 2W (Stockport, UK) with a 532 nm excitation. The filtered plasma samples (100 μl) were measured in a 4×4×10 mm optical cell with an antireflective coating (BioTools Inc., USA), which was placed in a home-made Peltier cell holder to control sample temperature (15 °C). The resolution of the ROA spectrometer is about 7 cm<sup>-1</sup>. The 4×4 mm quartz cell (BioTools, Inc., USA) was tempered to 15 °C. The filtered plasma (100 μl) was treated by 10 mg of NaI for the suppression of undesirable fluorescence background. The total exposure time for the ROA spectra was kept at 24 hours. Real laser power on the sample was monitored by an Optical power meter 1916-R with 818-P sensor (Newport, USA).

**Electronic circular dichroism.** The ECD spectra were measured on a J-815 spectrometer equipped with a Peltier unit CDF-426S/16 (Jasco, Japan) to control sample temperature. The filtered plasma was diluted four-times by phosphate buffer and measured in a 0.01 mm Suprasil quartz cuvette (Hellma, Germany) in the 280–185 nm spectral range. Six scans with a data pitch of 0.1 nm were accumulated for each sample and averaged in the Spectra Analysis module of the Spectra Manager program, ver. 2.6.0.1 (Jasco, Japan).

**Statistical data evaluation.** To discriminate between the patients with colorectal carcinoma and healthy controls, the obtained spectral data sets were evaluated by linear discriminant analysis (LDA) in the XLSTAT software (Addinsoft, France). A statistical model was created for selected spectral regions that include significant information about the structure of plasmatic proteins and other biomolecules. Sensitivity and specificity of the mathematical model were determined. The detrended Raman spectra and spectra from other methods (ECD, FT-IR, ROA) were processed using principal component analysis (PCA); and according to loadings, significantly contributing bands were selected. These bands were used for the discrimination of the control group and the group of patients by linear discriminant analysis (LDA). The further optimization of band selection was based on correlation coefficients, standardized canonical discriminant coefficients, and overall success in cross-validation.

## Results and discussion

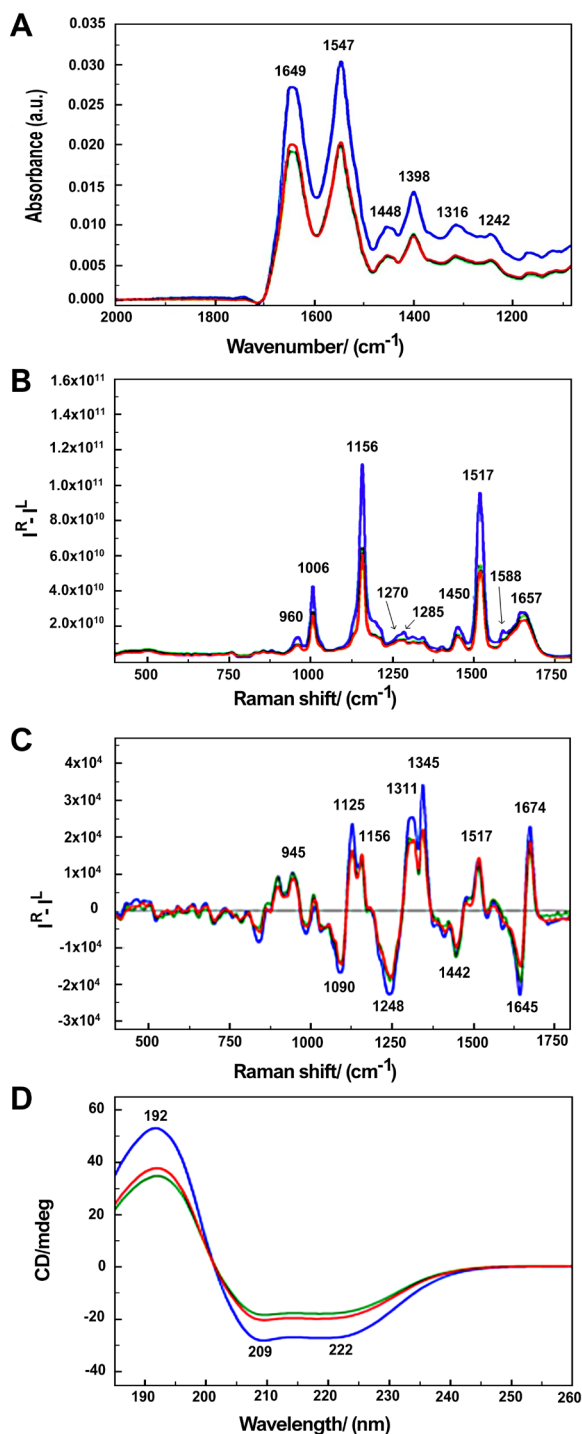
**Infrared spectroscopy.** The average FT-IR spectra are presented in Figure 1A. It shows two dominant bands with their maxima at 1649 and 1547 cm<sup>-1</sup> belonging to the vibra-

**Table 2. Patients and Tumor characteristics - Summary (n=30).**

Age	64.6 (37–86) years
Sex	
Male	20 (66.7%)
Female	10 (33.3%)
Tumor location	
Right colon	10 (33.3%)
Left colon	20 (66.7%)
Tumor morphology	
Adenocarcinoma grade 1	2 (6.7%)
Adenocarcinoma grade 2	21 (70.0%)
Adenocarcinoma grade 3	7 (23.3%)
Stage	
I	2 (6.7%)
II	9 (30.0%)
III	6 (20.0%)
IV	13 (43.3%)

Table 3. Selected biochemical parameters assessed from blood of patients with colorectal cancer and healthy controls.

	Albumin (g/l)	Total protein (g/l)	CRP (mg/l)	Glucose (mmol/l)	CEA ( $\mu$ g/l)	CA 19-9 (kIU/l)
Patient (moderate level)	21.4-45.6 (37.58)	46.8-74.1 (59.73)	1.0-223.9 (44.88)	4.7-19.5 (8.42)	0.5-353.4 (31.20)	2.1-9700.6 (685.74)
Control (moderate level)	44.7-55.5 (49.32)	65.1-86.0 (74.61)	1.0-20.4 (4.60)	2.7-7.4 (4.27)	0.8-7.1 (1.83)	2.0-20.1 (6.20)



tion of amide I and amide II, respectively. The intensity and shape of the  $1649 \text{ cm}^{-1}$  band localized in the amide I region are significantly influenced by the type and content of the protein secondary structure [22]. The second band at  $1547 \text{ cm}^{-1}$  arises from out-of-phase N-H in-plane bending and C $\alpha$ -N stretching of the peptide bond, the so-called amide II [22]. Lower intensities of both bands in the spectra of patients with colorectal carcinoma may again indicate variations in the secondary structure of plasmatic proteins. The band at  $1448 \text{ cm}^{-1}$  arose from the deformation vibrations of CH<sub>3</sub> and CH<sub>2</sub> groups in the side chains of proteins mixed with phospholipids. The  $1260\text{--}1390 \text{ cm}^{-1}$  spectral region coheres with the stretching vibration of protein carboxyl groups (phospholipids overlapping with cholesterol and COO<sup>-</sup> stretch of proteins) [23]. To compare both subgroups of patients, the bands of patients with early-stage CRC (stage I-II) are remarkably similar to patients with advanced-stage (III-IV).

**Raman spectroscopy.** The average Raman spectra (Figure 1B) show 3 pronounced bands corresponding primarily to the C-C and C=C vibrations of carotenoids ( $1006$ ,  $1156$ , and  $1517 \text{ cm}^{-1}$ ) [24-26]. Although carotenoids are only present in blood plasma at low concentrations [27], a resonance enhancement of their band intensities occurs with excitation in the visible spectral range ( $532 \text{ nm}$ ) [24, 28]. We observed significant changes within the carotenoid bands, in comparison with the control group, a significant intensity decrease of carotenoid bands was observed for the patients with CRC with all stages. Moreover, in some stage IV samples, exceptionally low intensity or even no carotenoid bands were observed. These findings suggest that carotenoids may play a role in the development and progress of CRC. Exhibiting strong antioxidant properties, carotenoids are apparently used to diminish the destructive impact of oxidative stress building up in the proximity to the growing tumor [27]. Other important bands at  $1657 \text{ cm}^{-1}$  (amide I),  $1270$ , and  $1285 \text{ cm}^{-1}$  (amide III) are related to proteins and vibrations of their peptide bonds. Spectral differences between the patients with CRC and healthy controls were detected

Figure 1. Average Fourier transform infrared spectra (A), Average Raman spectra (B), Average Raman optical activity spectra (C), and Average ECD spectra (D) of human blood plasma samples from patients with CRC stage I-II (green; n=11), patients with CRC stage III-IV (red; n=19) and healthy controls (blue; n=33).

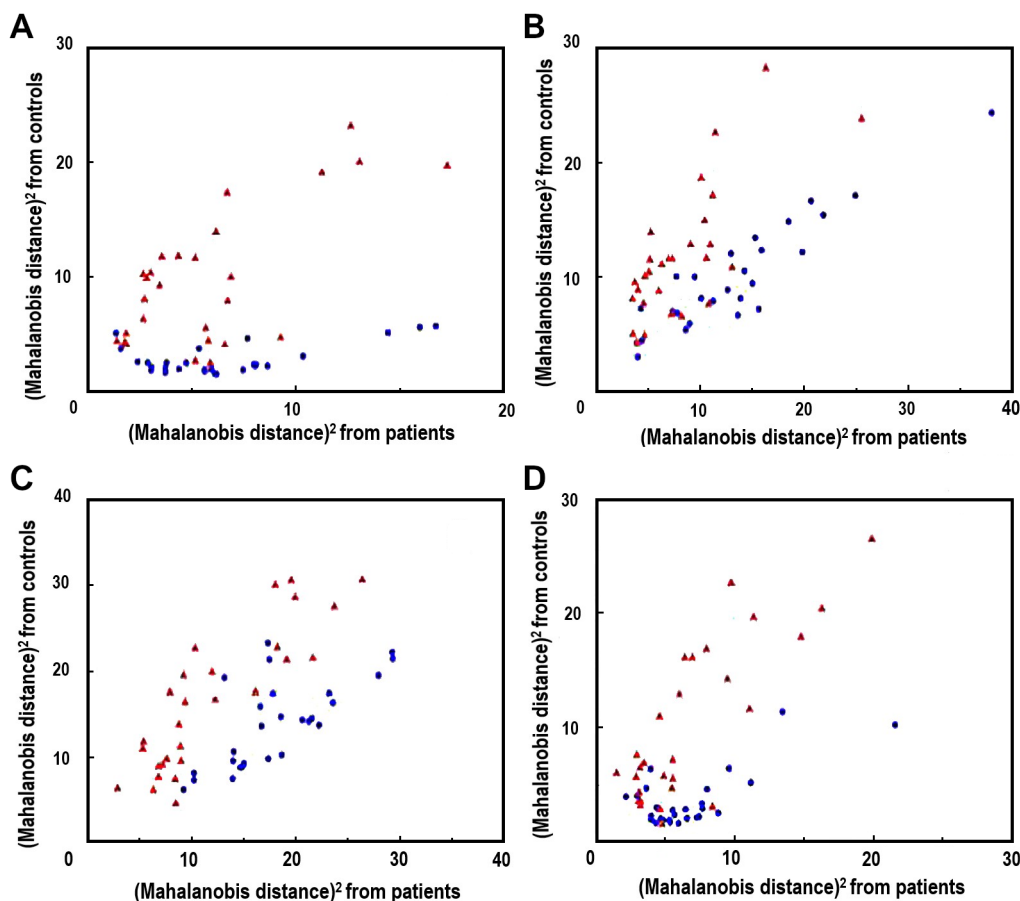


Figure 2. Graphical results of LDA for individual spectral methods, ECD (A), Raman (B), ROA (C), FT-IR (D); ● control group, ▲ patient group.

in the extended amide III region, arising from the combination of N-H and Ca-H bending vibrations of the peptide bond [29–31], the bands at this region were less intense for the patients than for controls. It indicates a major change in protein secondary structure and may point to partial protein misfolding and the loss of their correct biological function in the case of CRC. As expected, only minimal intensity and shape variations were observed at the 1657  $\text{cm}^{-1}$  band in the amide I region representing the C–O stretch of the peptide bond. Further spectral bands can be assigned to vibrations of glycoprotein moieties (960  $\text{cm}^{-1}$ ), side chain  $\text{CH}_3$  and  $\text{CH}_2$  groups overlapping with phospholipids (1450  $\text{cm}^{-1}$ ), tryptophan, and phenylalanine residues (1588, 1196, and 1006  $\text{cm}^{-1}$  overlapping with carotenoids) [29]. Changes within the majority of these bands most likely reflect the differences between the metabolic status of the patients with CRC and that of the healthy controls.

**Raman optical activity.** Figure 1C shows the average ROA spectra of blood plasma of the control and patient groups. There are several bands in the amide I (1630–1680  $\text{cm}^{-1}$ ; C–O stretch of the peptide bond) and extended amide III (1230–1350  $\text{cm}^{-1}$ ; in-phase combination of in-plane N–H

and Ca–H bends with a small contribution of Ca–N stretch of the peptide bond) regions [29–31]. The positive bands in the extended amide III region are especially sensitive to molecule geometry provided by the coupling between N–H and Ca–H vibrations [30]. Changes in the intensity and ratio of these bands can be also observed in this case. The negative-positive amide I couplet at  $-1645$  and  $+1674$   $\text{cm}^{-1}$  is characteristic for proteins with prevailing  $\alpha$ -helical conformation [30, 32, 33]. The intensity and/or pattern changes of other spectral bands were also observed, for example, within positive bands at 898 and 945  $\text{cm}^{-1}$  that confirm the prevailing  $\alpha$ -helical conformation of plasmatic proteins, a negative 1248  $\text{cm}^{-1}$  band reflecting more  $\beta$ -components or a negative band at 1442  $\text{cm}^{-1}$  indicating aliphatic side chains [30, 34]. The variations of these bands may demonstrate the disruption of serum albumin folding mechanism and/or structure. The prominent positive bands at 1156 and 1517  $\text{cm}^{-1}$  can be assigned to plasma carotenoids, which overlap with the bands of aromatic amino-acid residues [29].

**Electronic circular dichroism.** The average ECD spectra (Figure 1D) are dominated by a single positive band (192 nm) and 2 partially overlapping negative bands

**Table 4. Confusion matrices for the LDA cross-validation results showing an agreement between clinical diagnosis (rows) and spectroscopic analysis (columns) obtained from individual spectral methods.**

ECD spectroscopy					Raman spectroscopy				
From/to	Cancer	Control	Total	Correct (%)	From/to	Cancer	Control	Total	Correct (%)
Cancer	23	7	30	77	Cancer	21	9	30	70
Control	5	28	33	85	Control	9	24	33	73
Total	28	35	63	81	Total	30	33	63	71%

ROA spectroscopy					FT-IR spectroscopy				
From/to	Cancer	Control	Total	Correct (%)	From/to	Cancer	Control	Total	Correct (%)
Cancer	21	9	30	70	Cancer	22	8	30	74
Control	7	26	33	79	Control	6	27	33	81
Total	28	35	63	74	Total	28	35	63	77

(209, 222 nm), the slightly decreased intensities of which were observed in the case of the patients with CRC. These bands represent electronic transitions of the peptide bond in proteins [19], and their intensities and band shapes vary depending on the peptide backbone geometry [35, 36]. The overall spectral pattern is typical for alpha-helical proteins, due to ~60% content of albumin (34–47 g/l of albumin per 70–75 g/l of total proteins) in the blood plasma of healthy persons [15]. Thus, lower band intensities may correspond with partial albumin misfolding or the decrease in albumin level during the development of colorectal cancer. The main changes between the control and patient group are partly in the intensity of bands and partly in the relative ratio of the bands at 209 and 222 nm. The decreased intensity of these bands correlates with the biochemical data, the patient group had a lower ratio of albumin to total protein in the blood plasma than controls (Table 3).

**Statistical analysis.** To detect all differences included in the spectra, some of which could not be easily observed by the naked eye, the obtained data sets were processed by means of statistical evaluation. Linear discriminant analysis (LDA) was performed using spectral bands that reflect primarily protein secondary structure. For LDA, we combined 27 bands: 4 bands from FT-IR (1244, 1400, 1547, and 1639  $\text{cm}^{-1}$ ), 8 bands from Raman (1270, 1285, 1341, 1357, 1391, 1450, 1517, and 1586  $\text{cm}^{-1}$ ), 12 bands from ROA (833, 956, 1264, 1295, 1301, 1311, 1345, 1442, 1604, 1645, 1665, and 1674  $\text{cm}^{-1}$ ) and 3 bands from ECD (192, 209, and 222 nm). Most of these bands belong to proteins or other compounds described previously (e.g. carotenoids, phospholipids) in this article. The band selection was based on the evaluation of correlation and covariance matrix. The main aims of LDA were to differentiate the patients with CRC from healthy controls by spectral pattern recognition, to assess the sensitivity and specificity of the spectroscopic methods used, and to establish the reliability of our statistical model. The next aim was to differentiate the patients according to the stage of the CRC and by this method to determine the tumor load. Moreover, we evaluated spectra of 5 patients from the patient group, who have been undergone curative therapy and were 2 years with no signs of recurrence of the CRC

according to the clinical, laboratory, and imaging methods. We checked if their results were similar to healthy controls.

Because the highly structure-sensitive chiroptical spectroscopy provides additional information to the conventional vibrational spectroscopy, we created a statistical model that combines spectral data of all techniques used (FT-IR, Raman, ROA, and ECD). The spectra of each spectral method were statistically evaluated individually and compared (Figure 2). To emphasize the differences between individual samples and groups of samples, the LDA results are plotted in squared Mahalanobis distances [37]. Clearly, the control and the patient group were partly discriminated in each spectral method, but both groups overlapped.

The overall classification ability of models reached 81–87%. In addition, each spectral model was also processed by leave-one-out cross-validation (LOOCV), see Table 4.

According to LOOCV, ECD spectroscopy exhibited the highest sensitivity (77%) and specificity (85%), whereas the lowest sensitivity was observed in both ROA and Raman spectroscopy (70%). However, the lowest specificity (73%) was obtained from Raman spectroscopy, while in the case of ROA it reached 79%. The statistical evaluation of ECD data yielded the highest overall accuracy (81%) after LOOCV, whereas the lowest level (71%) was reached by Raman spectroscopy. Usually, the misclassified samples differed for each spectroscopic method. Each spectroscopy is sensitive to different properties of the studied biomolecules/samples. While ECD is sensitive to the electronic transitions of chromophores and their chirality, ROA is based on the change of the polarizability of functional groups and their chirality. Raman spectroscopy deals with the change of the polarizability of functional groups and FT-IR with the change of the dipole moment of functional groups.

Therefore, the obtained results were significantly improved by evaluating all spectral methods together, for compensating the inadequacies in the classification of certain samples by individual spectroscopic methods. The combined statistical model leads to significant classification improvement in comparison with individual techniques. After this combination, both the investigated groups were well separated (Figure 3A). The discrimination ability in

this case achieved 100%. The achieved results were verified LOOCV, where sensitivity and specificity reached 90% and 75%, respectively. 8 samples from the control group and 3 samples from the patient group were misclassified (Table 5). The overall accuracy for LOOCV was 83%.

As a next step in the spectral analysis, we divided the patients according to the clinical stages. The overall accuracy for LOOCV was 82% for the discrimination of the stage I+II, 64% for the discrimination of the stage III+IV, and 75% for the control group. Figure 3B illustrates the graphical result of LDA for the combination of FT-IR, Raman, ROA, and ECD spectroscopic data and separation of the control group, patients with clinical stages I–II and patients with clinical stages III–IV. As shown, there still were overlaps between different classifications. We have chosen 5 patients in clinical remission (“healthy patients”), which means, that they have undergone curative treatment and have got 2 years of no signs of cancer disease according to conventional clinical, laboratory, and imaging methods. The 4 patients were according to the spectroscopic methods evaluated as healthy controls, 1 patient (Patient 12 healthy) was assigned to a group of CRC patients with clinical stage III–IV. In this case, we have got suspicion on the cancer relapse also according to the imaging methods and we will check her early in the future. Figure 3C represents 3D graphical results of LDA for the combination of all spectral methods.

In conclusion, using a unique combination of chiroptical and vibrational spectroscopies for the analysis of blood plasma, we have successfully identified the spectral regions that are most likely affected by the pathological processes

of CRC. According to our observations, the most significant spectral differences between the blood plasma of CRC patients and healthy controls occur in spectral regions primarily reflecting the changes in protein secondary structure. Moreover, the very low intensities of carotenoid bands in the Raman spectra of CRC patients might be related to the disease progression and thus considered a spectral biomarker of CRC.

In this paper, we showed that chiroptical spectroscopy can be used in the combination with Raman and FT-IR spectroscopies for the discrimination of the control group from patients with CRC, moreover for the discrimination of the patients with CRC according to the tumor load. The detailed results of LDA showed that each spectral method (FT-IR, Raman, ROA, ECD) has limited ability to discriminate the control group from patients, but the best results were achieved for their combination. In this case, the discrimination ability achieved 100%. The mathematical model was sequentially confirmed by leave-one-out cross-validation, where sensitivity and specificity reached 90% and 85% respectively. The overall accuracy for the leave-one-out cross-validation was 83%. Regarding the discrimination among the control group vs. patients with limited-stage I–II and patients with advanced-stage III–IV (with regional lymphadenopathy or metastatic disease), the results were also promising. The overall accuracy was 82% for discrimination of the patients with CRC stage I–II, 64% for discrimination of the patients with CRC stage III–IV, and 75% for the control group. The achieved results indicate that chiroptical spectroscopy might be a new rapid supporting method in the clinical diagnosis of CRC based on the analysis of blood plasma, which is minimally invasive for patients and which also offers potentially early detection for CRC patients when there is no regional lymphadenopathy or metastatic disease presented. It might have a significant impact on reducing the incidence and improving the survival rates of CRC. Above that, this innovative method can predict cancer remission,

Table 5. Cross-validation results obtained from LDA of all spectral methods.

From/to	Cancer	Control	Total	Correct (%)
Cancer	27	3	30	90
Control	8	25	33	75
Total	35	28	63	83

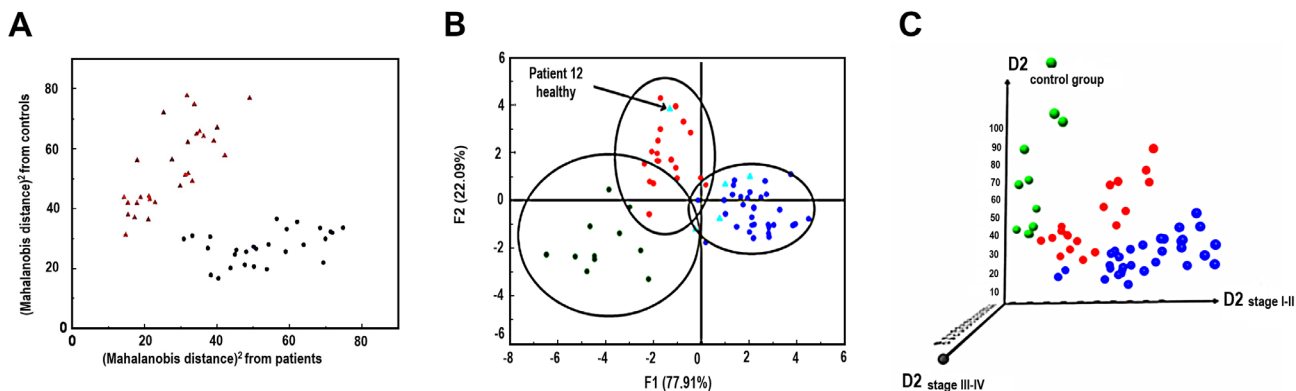


Figure 3. Graphical result of LDA for the combination of FT-IR, Raman, ROA, and ECD spectroscopic data A) in ● control group and ▲ patient group; B) in ● control group (blue), patient group with stage I–II (green), patient group with stage III–IV (red), healthy patients (turquoise); C) 3D graphical representation of the results of LDA for the combination of FT-IR, Raman, ROA, and ECD showing the differentiation between the controls (blue), patients with CRC stage I–II (green), patients with CRC stage III–IV (red).



according to our results with an accuracy of 80%. Our future work will focus on the sample set extension, data validation, and testing the interference with other diseases.

Acknowledgements: The work was financially supported by the Ministry of Health of the Czech Republic (project no. NU20-09-00229) and project PROGRES Q28/LF1.

## References

- [1] RAWLA P, SUNKARA T, BARSOUK A. Epidemiology of colorectal cancer: incidence, mortality, survival, and risk factors. *Prz Gastroenterol* 2019; 14: 89–103. <https://doi.org/10.5114/pg.2018.81072>
- [2] EWING I, HURLEY JJ, JOSEPHIDES E, MILLAR A. The molecular genetics of colorectal cancer. *Frontline Gastroenterol* 2014; 5: 26–30. <https://doi.org/10.1136/flgastro-2013-100329>
- [3] VOGELSTEIN B, FEARON ER, HAMILTON SR, KERN SE, PREISINGER AC et al. Genetic alterations during colorectal-tumor development. *N Engl J Med* 1988; 319: 525–532. <https://doi.org/10.1056/NEJM198809013190901>
- [4] DONG L, SUN X, CHAO Z, ZHANG S, ZHENG J et al. Evaluation of FTIR spectroscopy as diagnostic tool for colorectal cancer using spectral analysis. *Spectrochim Acta A Mol Biomol Spectrosc* 2015; 134: 638. <https://doi.org/10.1016/j.saa.2014.06.128>
- [5] SHORT M, TAI I, OWEN D, ZENG H. Using high frequency Raman spectra for colonic neoplasia detection. *Opt Express* 2013; 21: 5025–5034. <https://doi.org/10.1364/OE.21.005025>
- [6] FENG S, WANG W, TAI IT, CHEN G, CHEN R et al. Label-free surface-enhanced Raman spectroscopy for detection of colorectal cancer and precursor lesions using blood plasma. *Biomed Opt Express* 2015; 6: 3494–3502. <https://doi.org/10.1364/BOE.6.003494>
- [7] HARRIS AT, LUNGARI A, NEEDHAM CJ, SMITH SL, LONESMA et al. Potential for Raman spectroscopy to provide cancer screening using a peripheral blood sample. *Head Neck Oncol* 2009; 1: 34. <https://doi.org/10.1186/1758-3284-1-34>
- [8] REMBACKEN B, HASSAN C, RIEMANN JF, CHILTON A, RUTTER M et al. Quality in screening colonoscopy: position statement of the European Society of Gastrointestinal Endoscopy (ESGE). *Endoscopy* 2012; 44: 957–968. <https://doi.org/10.1055/s-0032-1325686>
- [9] ZHENG Q, KANG W, CHEN C, SHI X, YANG Y et al. Diagnosis accuracy of Raman spectroscopy in colorectal cancer: A PRISMA-compliant systematic review and meta-analysis. *Medicine (Baltimore)* 2019; 98: e16940. <https://doi.org/10.1097/MD.00000000000016940>
- [10] SATO S, SEKINE R, KAGOSHIMA H, KAZAMA K, KATO A et al. All-in-one Raman spectroscopy approach to diagnosis of colorectal cancer: analysis of spectra in the fingerprint regions. *J Anus Rectum Colon* 2019; 3: 84–90. <https://doi.org/10.23922/jarc.2018-039>
- [11] DAS NK, DAI Y, LIU P, HU C, TONG L et al. Raman plus X: Biomedical applications of multimodal Raman spectroscopy. *Sensors (Basel)* 2017; 17: 1592. <https://doi.org/10.3390/s17071592>
- [12] KRUCHININ VN, KRUCHININA MV, PRUDNIKOVA YI, SPESIVTSEV EV, RYKHLITSKII SV et al. The Use of Spectral Ellipsometry and Raman Spectroscopy in Screening Diagnostics of Colorectal Cancer. *Optics and Spectroscopy* 2019; 127(1): 170–176. <https://doi.org/10.1134/S0030400X19070142>
- [13] JENKINS CA, JENKINS RA, PRYSE MM, WELSBY KA, JITSUMURA M et al. A high-throughput serum Raman spectroscopy platform and methodology for colorectal cancer diagnostics. *Analyst* 2018; 143: 6014–6024. <https://doi.org/10.1039/c8an01323c>
- [14] JENKINS CA, LEWIS PD, DUNSTAN PR, HARRIS DA. Role of Raman spectroscopy and surface enhanced Raman spectroscopy in colorectal cancer. *World Journal of Gastrointestinal World J Gastrointest Oncol* 2016; 8: 427–438. <https://doi.org/10.4251/wjgo.v8.i5.427>
- [15] MURRAY RK, BENDER DA, BOTHAM KM, KENNELLY PJ, RODWELL VW, WEIL PA, editors. *Harper's Illustrated Biochemistry*. New York: McGraw-Hill, 2003.
- [16] LASCH P, KNEIPP J, editors. *Biomedical Vibrational Spectroscopy*. New York: John Wiley & Sons, Inc, 2008.
- [17] MITCHELL AL, GAJJAR KB, THEOPHILOU F, MARTIN FL, MARTIN-HIRSCH PL. Vibrational spectroscopy of biofluids for disease screening or diagnosis: translation from the laboratory to a clinical setting. *J Biophotonics* 2014; 7: 153–165. <https://doi.org/10.1002/jbio.201400018>
- [18] SCHIPPER HM, KWOK CS, ROSENDAHL SM, BANDILLA D, MAES O et al. Spectroscopy of human plasma for diagnosis of idiopathic Parkinson's disease. *Biomark Med* 2008; 2: 229–238. <https://doi.org/10.2217/17520363.2.3.229>
- [19] BEROVA N, NAKANISHI K, POLAVARAPU PL, WOODY RW, editors. *Comprehensive Chiroptical Spectroscopy: Applications in Stereochemical Analysis of Synthetic Compounds, Natural Products and Biomolecules*. Hoboken, New Jersey: John Wiley & Sons, Inc., 2012.
- [20] NAFIE LA, editor. *Vibrational Optical Activity*. Chichester: John Wiley & Sons, Inc., 2011.
- [21] MINDE DP, ANVARIAN Z, RUDIGER SG, MAURICE MM. Messing up disorder: how do missense mutations in the tumor suppressor protein APC lead to cancer? *Mol Cancer* 2011; 10: 101. <https://doi.org/10.1186/1476-4598-10-101>
- [22] DAIDONE I, ASCHI M, ZANETTI-POLZI L, DI NOLA A, AMADEI A. On the origin of IR spectral changes upon protein folding. *Chem Phys Lett*. 2010; 488: 213–218. <https://doi.org/10.1016/j.cplett.2010.02.020>
- [23] PIVA JAAC, SILVA JLR, RANIERO L, MARTIN AA, BOHR HG et al. Overview of the use of theory to understand infrared and Raman spectra and images of biomolecules: colorectal cancer as an example. *Theor Chem Acc*. 2011; 130: 1261–1273. <https://doi.org/10.1007/s00214-011-1063-0>
- [24] FELTL L, PACÁKOVÁ V, ŠTULÍK K, VOLKA K. Reliability of carotenoid analyses: a review. *Curr Anal Chem*. 2005; 1: 93–102. <https://doi.org/10.2174/1573411052948424>
- [25] WITHNALL R, CHOWDHRY BZ, SILVER J, EDWARDS HGM, DE OLIVEIRA LFC. Raman spectra of carotenoids in natural products. *Spectrochim Acta A Mol Biomol Spectrosc* 2003; 59: 2207–2212. [https://doi.org/10.1016/s1386-1425\(03\)00064-7](https://doi.org/10.1016/s1386-1425(03)00064-7)

- [26] PARKER SF, TAVENDER SM, DIXON NM, HERMAN HK, WILLIAMS KPJ et al. Raman spectrum of beta-carotene using laser lines from green (514.5 nm) to near-infrared (1064 nm): implications for the characterization of conjugated polyenes. *Appl Spectrosc.* 1999; 53: 86–91.
- [27] PARKER RS. Carotenoids in human blood and tissues. *J Nutr* 1989; 119: 101–104. <https://doi.org/10.1093/jn/119.1.101>
- [28] TATARKOVIČ M, SYNITSYA A, ŠŤOVÍČKOVÁ L, BUNGANIČ B, MIŠKOVIČOVÁ M et al. The minimizing of fluorescence background in Raman optical activity and Raman spectra of human blood plasma. *Anal Bioanal Chem* 2015; 407: 1335–1342. <https://doi.org/10.1007/s00216-014-8358-7>
- [29] SYNITSYA A, JUDEXOVÁ M, HRUBÝ T, TATARKOVIČ M, MIŠKOVIČOVÁ M et al. Analysis of human blood plasma and hen egg white by chiroptical spectroscopic methods (ECD, VCD, ROA). *Anal Bioanal Chem* 2013; 405: 5441–5453. <https://doi.org/10.1007/s00216-013-6946-6>
- [30] BARRON LD, ZHU F, HECHT L, TRANTER GE, ISAACS NW. Raman optical activity: an incisive probe of molecular chirality and biomolecular structure. *J Mol Struct.* 2007; 834–836: 7–16. <https://doi.org/10.1016/j.molstruc.2006.10.033>
- [31] ZHU F, ISAACS NW, HECHT L, BARRON LD. Raman optical activity: a tool for protein structure analysis. *Structure* 2005; 13: 1409–1419. <https://doi.org/10.1016/j.str.2005.07.009>
- [32] KINALWA MN, BLANCH EW, DOIG AJ. Accurate determination of protein secondary structure content from Raman and Raman optical activity spectra. *Anal Chem* 2010; 82: 6347–6349. <https://doi.org/10.1021/ac101334h>
- [33] URBANOVÁ M. Bioinspired interactions studied by vibrational circular dichroism. *Chirality* 2009; 21: E215–230. <https://doi.org/10.1002/chir.20803>
- [34] WEYMUTH T, REIHER M. Characteristic Raman optical activity signatures of protein  $\beta$ -sheets. *J Phys Chem B* 2013; 117: 11943–11953. <https://doi.org/10.1021/jp405981h>
- [35] WHITMORE L, WALLACE BA. Protein secondary structure analyses from circular dichroism spectroscopy: methods and reference databases. *Biopolymers* 2008; 89: 392–400. <https://doi.org/10.1002/bip.20853>
- [36] KELLY SM, JESS TJ, PRICE NC. How to study proteins by circular dichroism. *Biochim Biophys Acta* 2005; 1751: 119–139. <https://doi.org/10.1016/j.bbapap.2005.06.005>
- [37] DE MAESSCHALCK R, JOUAN-RIMBAUD D, MASSART DL. The Mahalanobis distance. *Chemometrics Intel Lab Syst.* 2000; 50: 1–18. [https://doi.org/10.1016/S0169-7439\(99\)00047-7](https://doi.org/10.1016/S0169-7439(99)00047-7)