analytical chemistry

pubs.acs.org/ac
Article

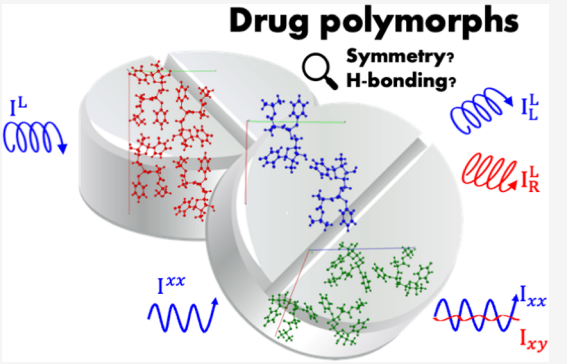
Sofosbuvir Polymorphs Distinguished by Linearly and Circularly Polarized Raman Microscopy

Věra Schrenková, Josef Kapitán, Petr Bouř, Argyro Chatziadi, Adam Sklenář, and Jakub Kaminský*

Cite This: *Anal. Chem.* 2024, 96, 18983−18993
Read Online

ACCESS | Metrics & More | Article Recommendations | Supporting Information

ABSTRACT: Most currently marketed pharmaceuticals are manufactured in the solid state, where the bioavailability of the active pharmaceutical ingredient (API) can be optimized through different polymorphs, cocrystals, solvates, or salts. Efficient techniques are needed to monitor the structure of pharmaceuticals during production. Here, we explore the potential of linearly and circularly polarized Raman microscopy for distinguishing three polymorphs of sofosbuvir, an antiviral drug used to treat hepatitis C. Raman spectra were recorded on a Raman microscope for a polycrystalline API diluted in a KBr matrix. To understand spectral features including the low-frequency region, we simulated band frequencies and intensities using quantum-chemical computational strategies based on cluster and transfer approaches. Very good agreement was achieved between simulated and experimental intensities. The 20 to 200 cm^{-1} wavenumber region appeared particularly useful for polymorph discrimination already based on unpolarized measurements. The depolarization ratios obtained from linearly polarized Raman spectra made the distinction even more reliable. Moreover, circularly polarized Raman spectra and normalized degrees of circularity provided useful additional information and revealed several tentative markers of the different polymorphs of sofosbuvir. Although in some spectral regions the differences were less obvious, the results indicate that polarized Raman microscopy is a handy tool for discriminating between polymorphs of APIs and other compounds.



INTRODUCTION

Most marketed drugs are formulated in the solid state, but their poor water solubility often limits their bioavailability and therapeutic efficiency.[1] Formulations such as cocrystals,[2] salts,[3] solvates[4] or polymorphs,[5] all allow fine-tuning of physical and chemical properties of an active pharmaceutical ingredient (API).[6−8] Crystal polymorphism, where a compound can adopt multiple crystal structures, is reported for more than half of solid drugs,[5,9] reliable and reasonably fast analytical methods for structural monitoring are needed to ensure proper product functionality.

Most often, diffraction techniques are used,[10] followed by solid-state nuclear magnetic resonance[11,12] and thermal methods.[13] Vibrational techniques such as Raman[14−16] and infrared[17] spectroscopy offer a promising alternative due to their nondestructive nature and sensitivity to differences in crystalline arrangement. A chiral variant of infrared absorption, vibrational circular dichroism, has also proven capable of distinguishing between polymorphs.[18−20] Also, coupling Raman spectroscopy with confocal microscopy appears convenient because it makes it possible to visualize the distribution of APIs within tablets,[21] monitoring drug release,[22] and tracking changes during pharmaceutical production.[23]

With the recent availability of efficient optical filters, the low-frequency (LF) region (sometimes referred to as the terahertz region) has become commonly available on commercial spectrometers and Raman microscopes.[24,25] LF Raman bands (<200 cm^{-1}) are particularly sensitive to intermolecular interactions and therefore closely reflect crystal packing.[26] It has been shown that LF Raman spectra effectively distinguish drug polymorphs,[15,27,28] or racemic crystals from enantiopure crystals,[29] cocrystals[30] and pharmaceutical excipients.[31]

Compared to unpolarized Raman experiments, additional information on the symmetry of vibrational modes and orientation of crystalline materials can be obtained by controlling the polarization states of incident and scattered light.[32−34] Although recent advances allow for the simultaneous acquisition of all polarization states,[35] in our present study we performed separate polarization measurements at each scanning point.

Still, there are relatively few studies that use polarized Raman spectroscopy methods, especially in pharmaceutical applications. In linearly polarized Raman spectroscopy (LPRS), the incident light is linearly polarized, and a parallel

Received: July 10, 2024
Revised: October 26, 2024
Accepted: November 13, 2024
Published: November 21, 2024

or perpendicular component of the scattered light passes through the analyzer. LPRS has been used to study the crystallographic orientation of single crystals,[36] 2D materials[37] and nanowires.[38] Also, LPRS can be used to ascertain the preferential orientation of API particles on the surface of tablets[35,39] or different intermolecular hydrogen bonding over a wide temperature range in paracetamol polymorphs.[40]

Similarly, in circularly polarized Raman spectroscopy (CPRS), incident and scattered light is circularly polarized, either in the same or opposite helicity.[41] CPRS has been applied, in particular, to investigate lattice properties of layered materials[42] and to determine the axial angle in chiral nanotubes.[43] As far as we know, no application of CPRS to solid-phase pharmaceutical samples has hitherto been reported.

Strictly speaking, CPRS does not provide additional information compared to LPRS, as the same molecular properties are involved.[44] However, the practical advantages of the two approaches remain to be ascertained. For example, differences between polymorphs may be better visible in CPRS, or experimental artifacts might be smaller with CPRS than with LPRS. Either polarization technique can enhance spectral resolution by resolving bands that overlap in nonpolarized experiments.[45]

Alongside experimental advancements, computational techniques have greatly improved the interpretation of both polarized and unpolarized Raman spectra in both solutions and solids.[46−49] Plane−wave (PW) density functional theory (DFT) is often applied to crystals, as demonstrated for ribavirin polymorphs.[50] Here, we use PWDFT[51] only to improve X-ray geometries, because the method appears to be too slow for studying the vibrational properties of large molecules. Therefore, we combine or replace PW computations with cluster-based and Gaussian basis set methods. Because of the size of sofosbuvir, we use Cartesian coordinate tensor transfer (CCT)[52] to calculate important intermolecular interactions taking place within crystals. Earlier, this approach was used to calculate Raman spectra of crystalline and amorphous poly(3-hydroxybutyrate),[53] as well as Raman spectra of methacrylamide, piracetam and 2-thiobarbituric acid polymorphs.[54]

Below, we present Raman spectra (unpolarized, LPRS and CPRS) of three polymorphs of the antiviral drug sofosbuvir across a broad spectral range (20−3500 cm^{-1}). Sofosbuvir (Figure 1) is used to treat chronic hepatitis C (HCV) because

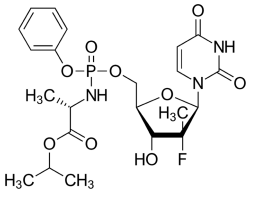


Figure 1. Structure of sofosbuvir.

it inhibits NS5B nucleotide polymerase.[55] Sofosbuvir's flexibility related to the phosphate group contributes to its rich polymorphism. At least 15 polymorphs of sofosbuvir have been reported so far,[56] along with its cocrystals or solvates.[57] The spectra of the polymorphs studied here are recorded for a polycrystalline API pressed into pellets, eliminating the need for single crystals. DFT simulations then allow us to more accurately identify relationships between the structure of compounds and their spectra. This study demonstrates the potential of polarized Raman microscopy for the precise characterization of pharmaceutical polymorphs, providing a valuable tool for drug development.

METHODS

Sofosbuvir Structure. Polymorphs chosen for this study are denoted as form 1, form 6 and form 7 (Figure 2), and they were prepared by Chatziadi et al.[58] Form 1 crystallizes in a monoclinic crystal lattice with the $P2_1$ space group and comprises four molecules in the elementary cell. Form 7 also crystallizes in a monoclinic lattice (space group $P2_1$), but it contains two molecules in the elementary cell.

Form 6 is the most compact and stable at room temperature, adopting an orthorhombic crystal lattice with the $P2_12_12_1$ space group, and there are four molecules in the elementary cell. The conformation of sofosbuvir in form 1 differs significantly from that in the others. Forms 6 and 7 are conformationally similar, as they differ mainly in the orientation of the L-alanine isopropyl ester group (marked as C in Figure 2). Forms 6 and 7 create hydrogen-bonded chains that link similar sofosbuvir rotamers. In form 7, the hydrogen bond chains are oriented in the same direction whereas in form 6 they alternate (i.e., are flipped by 180°). All three forms have been recently studied using vibrational circular dichroism.[20]

LPRS and CPRS Parameters. In LPRS, the ratio of perpendicular ($I_\perp$) to parallel ($I_\parallel$) intensities with respect to the polarization plane of the incident light is called the depolarization ratio $\rho = I_\perp/I_\parallel$. For isotropic samples it can be related to molecular parameters as

$$\rho = \frac{6\beta(\alpha^2)}{45\alpha^2 + 7\beta(\alpha^2) + [456\alpha^2 + \beta(\alpha^2)]P\cos 2\eta\cos 2\theta}$$

where α^2 and $\beta(\alpha^2)$ are invariants of the polarizability tensor and P, θ and η describe the polarization and geometry of the experiment. Parameter ρ reflects the symmetry of vibrational modes; totally symmetric modes give $\beta(\alpha^2) = 0$ and $\rho = 0$ whereas asymmetric modes give $\alpha^2 = 0$ and ρ can reach 3/4, the maximum value for a backscattering experiment.[25] The value of ρ is generally nonzero.

In analogy to ρ, the degree of circularity (DOC) is defined in CPRS as

$$\mathrm{DOC(R)} = \frac{I_R^R - I_L^R}{I_R^R + I_L^R} \wedge \mathrm{DOC(L)} = \frac{I_R^L - I_L^L}{I_R^L + I_L^L}$$

where the upper and the lower index denotes the helicity of the incident or scattered light, respectively. For a backward scattering geometry (180°), it can be calculated from as

$$\mathrm{DOC} = \frac{-5P\sin 2\eta[9\alpha^2 - \beta(\alpha^2)]}{45\alpha^2 + 7\beta(\alpha^2)}$$

where $P = 1$ and $\eta = \pm\frac{\pi}{4}$ for right- or left-handed circularly polarized incident light. The scattered light completely reverses circular polarization for $\beta(\alpha^2) = 0$ (DOC = −1), and it becomes partially circularly polarized (DOC = $\frac{5}{7}$) in the same sense if $\alpha^2 = 0$; DOC of fluorescence signal is close to zero (apart from fluorescence-detected CD signal, CPL, etc.). Under backscattering conditions, one can convert DOC to ρ, for example to verify experimental CPRS results, as

https://doi.org/10.1021/acs.analchem.4c03573
Anal. Chem. 2024, 96, 18983−18993

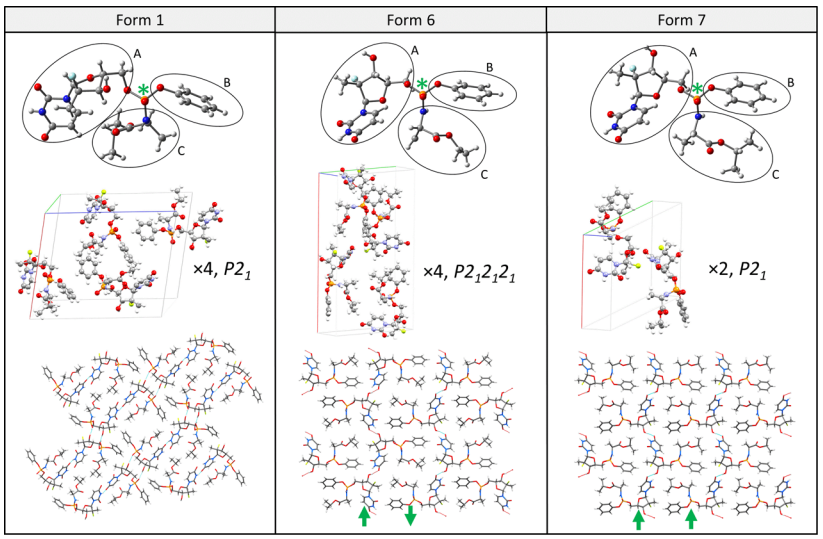


Figure 2. Single molecules, elementary cells and larger parts with H-bonded chains of sofosbuvir polymorphs under study,[58] denoted as form 1, form 6 and form 7. A = nucleotide group, B = phenyl group, C = L-alanine isopropylester group. The phosphorus atom is marked with an asterisk (*) for better orientation.

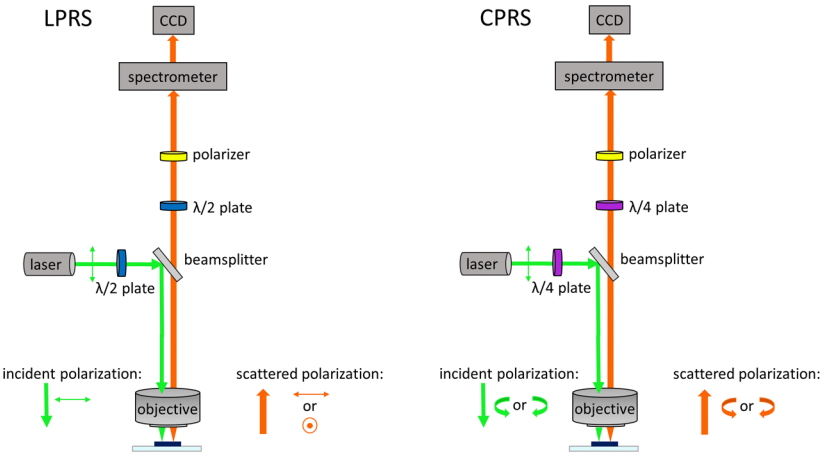


Figure 3. Raman microscope setup for linearly (left) and circularly (right) polarized modes. Unpolarized (I_u), parallel ($I_\parallel$) and perpendicular ($I_\perp$) spectra were recorded in the LPRS mode, corotating (I_L^L, I_R^R) and contrarotating (I_R^L, I_L^R) spectra in the CPRS mode.

$$\rho = \frac{1 + \mathrm{DOC(R)}}{3 - \mathrm{DOC(R)}}$$

Sample Preparation. Three polycrystalline polymorphs of sofosbuvir (form 1, form 6 and form 7) were prepared according to previously reported protocols.[58] The purity of the polymorphs was confirmed by measuring their melting point temperatures (form 1 = 96 °C, form 6 = 118 °C, form 7 = 125 °C), which were consistent with previously published values.[20,58] Raman maps were recorded for polycrystalline sofosbuvir in KBr. This matrix is regularly used in infrared spectroscopy and appeared convenient also for Raman experiments. It enabled us to homogenize samples and reduce polarization artifacts. Also, KBr to some degree mimics the environment within medicinal tablets. Sofosbuvir and potassium bromide (Sigma-Aldrich, ≥97%) were finely ground together in an agate mortar in a 1:1 ratio. The resulting mixture was compressed into a thin transparent pellet (7 mm in diameter) using a laboratory hydraulic press (Specac Mini Pellet Press). Following ref 20, polymorphic purity was additionally checked by *k*-means cluster analysis of the spectra obtained from the maps, which confirmed the occurrence of only one form in the sample.

Data Acquisition. Raman experiments were carried out using two setups of the WITec Alpha 300RS confocal Raman microscope (Figure 3).

We recorded a spectral map for each polymorph in each optical arrangement (parallel and perpendicular setup for LPRS, corotating and contrarotating setup for CPRS). For LPRS, the microscope was equipped with two half-wave plates placed in front and behind the sample, and the light was analyzed using a polarizer. For CPRS, we used a quarter-wave plate to convert linearly polarized light coming from the laser into circularly polarized light. The scattered light then passed through a second quarter-wave plate and the polarizer. Raman mapping of the pellets was performed in a backscattering geometry. A 10× dry objective lens (Zeiss EC Epiplan, numerical aperture 0.25) was used to focus the laser beam onto the sample and to collect the scattered light. A laser with a wavelength of 532 nm and a power of 19 mW, with 1800 g/mm diffraction grating for spectra below 200 cm^{-1} and 600 g/mm grating for spectra at and above 200 cm^{-1}, was used. The

https://doi.org/10.1021/acs.analchem.4c03573
Anal. Chem. 2024, 96, 18983−18993

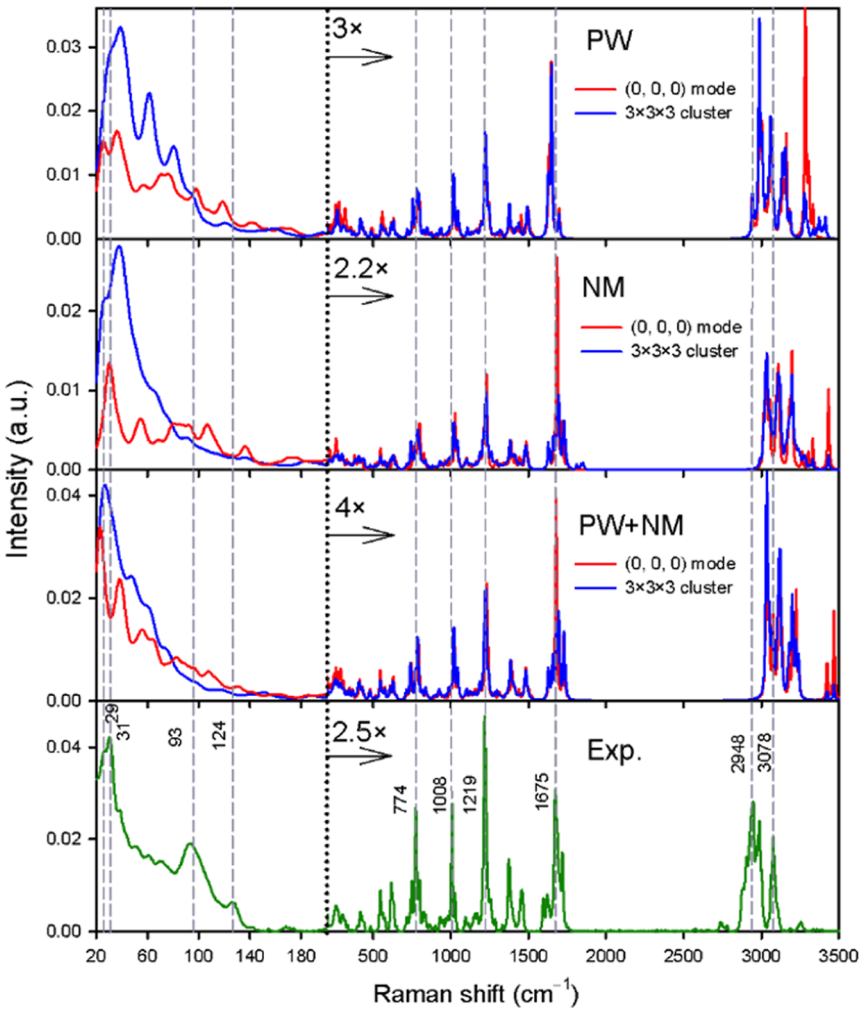


Figure 4. Comparison between calculated unpolarized Raman spectra of sofosbuvir form 7 and experimental data. The B3LYP/6-311++G(d,p)/CPCM level was used. Frequencies are not scaled and integral intensity is normalized against experimental values within the 200−1800 cm^{-1} range.

Raman mapping was done for an area of 300 × 300 at 2 μm steps (150 × 150 points) with 1 s integration time. Each image consisted of 22,500 spectra.

Data Processing. Raman maps were preprocessed using WITec Project 5.1 software. First, we applied a cosmic ray removal filter (filter size 4, dynamic factor 5) and then we averaged the signal over all pixels. Because Raman bands are narrower than other signal components, the baseline was subtracted from the spectra using our own implementation of the rubberband algorithm.[59]

Spectral Simulations. Spectra of all sofosbuvir polymorphs were calculated at the DFT level for structures derived from X-ray geometries (CCDC entries CUZROG01, CUZROG02 and CUZROG03) using the fragmentation and tensor transfer approach, which has been successfully used to model the vibrational spectra of various compounds in the condensed[20] or semicondensed state.[60] All details of the procedures can be found in the Supporting Information.

Because the positions of hydrogen atoms determined by X-ray diffraction may not be exact, we used three computational strategies to refine our geometries. In the first approach, we optimized the geometry of the unit cell using the PW basis sets technique as implemented in the program CASTEP[51] before assembling the 3 × 3 × 3 fragment. Alternatively, we partially optimized only the molecular pairs detected and created from the fragment, using the QGRAD program[61] interfaced to Gaussian 16 (ref 62) in normal mode (NM) vibrational coordinates.[63] In the third, "combined" approach, we preoptimized the crystal unit cell in CASTEP and then optimized the molecular pairs using the NM coordinates.

The harmonic force field and polarizability derivatives were then calculated for the optimized molecular pairs, using the program Gaussian[62] at the B3LYP/6-311++G(d,p) level. The crystal environment was simulated using the CPCM solvent model[64] and relative permittivity ε_r = 78.[54] Subsequently, the atomic property tensors were transferred back to the 3 × 3 × 3 fragment, using the CCT scheme.[52] Vibrational frequencies and Raman intensities were calculated in two ways: (1) directly for the 3 × 3 × 3 cluster and (2) using a dynamic matrix constructed from the same force field and the zero (0, 0, 0) crystal phonon mode. Line intensities, depolarization and DOC ratios were convoluted with Lorentzian bands of 10 cm^{-1} full width at half height. Frequencies are not scaled, and integral intensities were normalized against experimental values within the 200−1800 cm^{-1} range.

RESULTS AND DISCUSSION

Theoretical Models. Figure 4 shows the effect of geometry optimization and vibrational modeling on calculated unpolarized Raman spectra of form 7, in comparison with

https://doi.org/10.1021/acs.analchem.4c03573
Anal. Chem. 2024, 96, 18983−18993

Table 1. Experimental and Calculated Frequencies of Selected Raman Bands of Sofosbuvir Form 7

	PW optimization		NM optimization		PW + NM optimization	
exp.	3 × 3 × 3 crystal	dynamic cell	3 × 3 × 3 crystal	dynamic cell	3 × 3 × 3 crystal	dynamic cell
29	30	26	26		27	23
31	39	36	38	30		38
	61	71		54	47	56
93	80	98		81		82
124		119		107		
774	780	786	784	799	786	792
1008	1018	1019	1019	1027	1020	1022
1219	1222	1220	1232	1230	1224	1231
1675	1649	1648	1694	1649	1695	1678
2948	2988	2988	3042	3036	3036	3040
3078	3278	3284	3195	3198	3201	3223

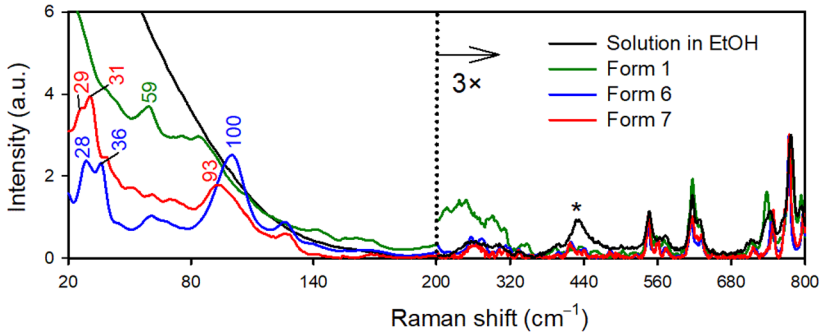


Figure 5. Unpolarized experimental Raman spectra of solid state sofosbuvir forms 1, 6, and 7 and of a 100 mg/mL solution of sofosbuvir dissolved in ethanol (the asterisk marks the ethanol band).

experimental data. The optimization strategies included the PW, NM optimization and a combined approach (see the Methods section). The calculated and experimental frequencies of the strongest bands are compared in more detail in Figure SI and Table 1.

Although the main spectral features are well reproduced by all three optimization approaches, in the lowest-wavenumber region the spectral shapes strongly depend on fine differences in the optimization. PW optimization inherently includes long-range interactions in the crystal. However, employing a different DFT level, a different description of a basis set for the subsequent calculation of vibrational properties may introduce errors. On the other hand, NM optimization is performed at the same theoretical level as the force field, but only molecular pairs without long-range interactions are considered in the calculation. The calculated frequencies correspond closely to equilibria on the potential energy surface, particularly for modes that were not constrained. The combined approach may compensate for some of the disadvantages of the two approaches.

The bands within the ~20 to 200 cm^{-1} range result from modes involving translations/rotations of various molecular parts and whole crystal layers. This spectral part is better reproduced by the zero-crystal phonon model, which otherwise gives results almost indistinguishable from those obtained for the 3 × 3 × 3 cluster. The "3 × 3 × 3" spectra below 200 cm^{-1} are less structured than in experimental data. Supposedly, the bands at 30, 93, and 124 cm^{-1}, calculated using the PW method, correspond to experimental signals at 36, 98, and 119 cm^{-1}. All calculations predict a large intensity below 200 cm^{-1}.

In the fingerprint region (~200 to 1800 cm^{-1}), the PW method was in slightly worse agreement with experimental data than the others. This is apparent for the C═N stretching and NH_2 bending vibrational pattern, experimentally determined to occur at around 1675 cm^{-1}, where the PW method produces an unrealistically sharp band at 1649 cm^{-1}. However, given its good performance with LF Raman spectra, this is only a minor drawback of PW optimization, which can still be recommended as a universal approach to first-principles calculations.

Within the 2800−3500 cm^{-1} range, all calculated frequencies were overestimated, mostly because the anharmonic effects were neglected.[54,65] The relative intensity of out-of-phase and in-phase NH bands (>3200 cm^{-1}) is often overestimated by calculations.

Recognition of Polymorphs in the LF Region. Whereas only minor differences among the polymorphs were observed within 200−3500 cm^{-1} range (Figure SII), larger variations occur below 200 cm^{-1} (Figure 5). In Figure 5, we also included the spectrum of sofosbuvir dissolved in ethanol, to document the better-resolved bands in the solid state, which are sensitive to crystal packing.[66]

The differences in the spectra of polymorphs in the LF region partially reflect the conformation of sofosbuvir but are influenced to a greater degree by the environment within the crystal and by crystal packing. To demonstrate that, we compared spectra of the conformationally similar polymorphs 6 and 7 calculated for isolated molecules in a vacuum with those obtained for a crystal (see Figure SIII). We can see that the spectra of isolated molecules are similar whereas those of molecule in different crystal arrangements differ much more.

In form 7, the H-bonded chains are oriented in the same direction whereas in form 6 they alternate. The LF Raman spectrum of form 1 notably differs from those of form 6 and 7

https://doi.org/10.1021/acs.analchem.4c03573
Anal. Chem. 2024, 96, 18983−18993

(see also Figure SII), likely reflecting the difference in conformation. Crystal packing does not seem so important. The same trend was previously observed for chlorpropamide, forms II and IV of which exhibited similar LF Raman spectra despite having different crystal packing due to the orientation of the aromatic ring between the consecutive molecules.[67] On the other hand, forms III and IV of chlorpropamide had distinct LF Raman bands owing to conformational polymorphism and the head-to-tail orientation of chlorpropamide molecules.[67] For carbamazepine polymorphs, it has been proposed that crystals having greater density and higher symmetry produce bands with lower frequencies.[68] However, no such trend was observed for theophylline polymorphs.[68] Similarly, for sofosbuvir polymorphs, the density of which decreases from form 6 to form 7 and form 1 (6 > 7 > 1) and whose symmetry is higher in form 6 than in forms 1 and 7 (6 > 1, 7), we observe no correlation between density, symmetry and band frequency. For a better idea, visualizations of the computed LF modes (PW, zero phonon mode) are presented in Figure SIV.

The intensity of LF bands is approximately three times greater than that of typical fingerprint bands. This may be explained by the large amplitude of LF vibrations, their temperature-induced excitations and the movement of highly polarizable π-systems.[68,69] To estimate the contributions originating from the phenyl group and the uracil ring, we set the polarizability derivatives with respect to other atoms to zero. The resulting spectra are plotted in Figure 6. Indeed, we see that the phenyl signal dominates in the LF region. This means that the strongest band at 79 cm^{-1} involves phenyl motion and dimethyl deformation with contributions from the uracil ring and deformation of the methyls in the L-alanine isopropylester group. Of course, the rings, especially the of uracil, also significantly contribute in the fingerprint region.

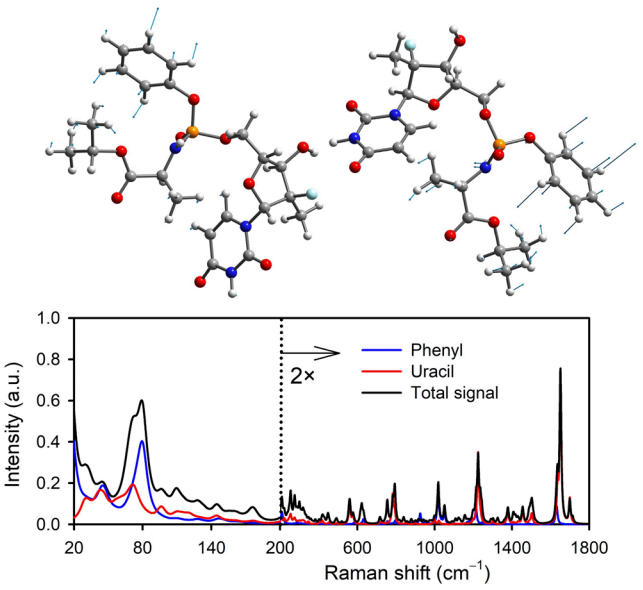


Figure 6. Examples of LF NMs of form 6 of sofosbuvir and the approximate contributions of aromatic rings to the total Raman signal.

Linearly Polarized Raman Spectra. Calculated (PW optimization, zero phonon mode) and experimental LPRS spectra of sofosbuvir forms 1, 6, and 7 are provided in Figure 7. The assignment of selected bands and depolarization ratios for the different computational models used are listed in Tables SI–SIV.

In Figure 7, we observe good agreement between calculated and experimental data also for the LPRS spectra. In contrast to the fingerprint region, where the depolarization ratios are rather low, in the LF region these $\rho_{Exp.}$ are greater than 0.50. This may reflect the contribution of intermolecular vibrations, which are more complex and more asymmetrical than localized vibrational modes.[26] Raman signal in the fingerprint region (~200 to 1800 cm^{-1}) involves relatively localized vibrations. The bands are therefore generally less sensitive to the conformation of the molecule and to the environment than LF vibrations. This does not apply to C=O or O–H stretching, which is sensitive to the environmental interactions mediated by hydrogen bonding. Although some differences between the three crystalline forms are observable, they are not as pronounced as in the LF region. In spite of the different crystal packing in forms 6 and 7, their LPRS in the fingerprint region is very close. For example, both form 6 and form 7 have an intense band corresponding to C=N and C=O valence stretching vibrations mixed with N–H deformation found at 1674 cm^{-1} (form 6) or at 1675 cm^{-1} (form 7). For form 1, this peak is shifted to 1685 cm^{-1} and is broader. Analogous similarity between different polymorphs with different packing, but in unpolarized Raman spectra, has been described also for piracetam.[54]

The hydrogen-stretching region (>2800 cm^{-1}) appears to be the least convenient for discriminating between the polymorphs, although some differences can be found in this region as well. For the symmetric methyl stretching at ~2944 cm^{-1}, depolarization varies the most (form 1—$\rho_{Exp.}$ = 0.08, form 6—$\rho_{Exp.}$ = 0.16, form 7—$\rho_{Exp.}$ = 0.10). For form 1, the aryl signal occurs at 3074 cm^{-1} whereas it is shifted to 3078 cm^{-1} in form 6 and 3079 cm^{-1} in form 7.

Because the differences in the fingerprint region and the hydrogen-stretching region (>200 cm^{-1}) are less apparent, we analyze them in more detail in Figure 8 (panel a). Form 6 exhibits the highest values of $\rho_{Exp.}$, and forms 1 and 7 are similar. Figure 8, panel b summarizes the distributions of $\rho_{Exp.}$; the average of form 7 (0.15) is identical to that of form 1, the range is the widest for form 7, and form 6 has the highest average (0.22) and the narrowest range. Nevertheless, the correlation coefficients (Figure 8, panel c) are very high across the polymorph pairs.

The calculated depolarization ratios (Figure 9) well correlate with the experimental ones, although the calculations often do not reproduce the fine differences between the polymorphs, presumably because of band overlap and a lack of anharmonic interactions (for the C–H stretching) in the modeling.

Circularly Polarized Raman Spectra. Our experimental results indicate that it is possible to differentiate polymorphs based on their DOC ratios (Figure 10). For example, the DOC band at 322 cm^{-1} of form 1 is approximately three times more intense than those of form 6 and form 7. Similarly, the DOC(L) band at 455 cm^{-1} is positive for form 1 whereas for form 6 and form 7 the DOC is zero. Form 6 has the lowest values of DOC among the three forms. This correlates with the depolarization ratios, as large DOCs correspond to low ρ and form 6 has the highest ρ values. In Table SV we compare DOC(R) values calculated from $\rho_{Exp.}$ for six bands with experimental DOC(R) values, which indicates a good correlation with an error of 0–4%.

https://doi.org/10.1021/acs.analchem.4c03573
Anal. Chem. 2024, 96, 18983–18993

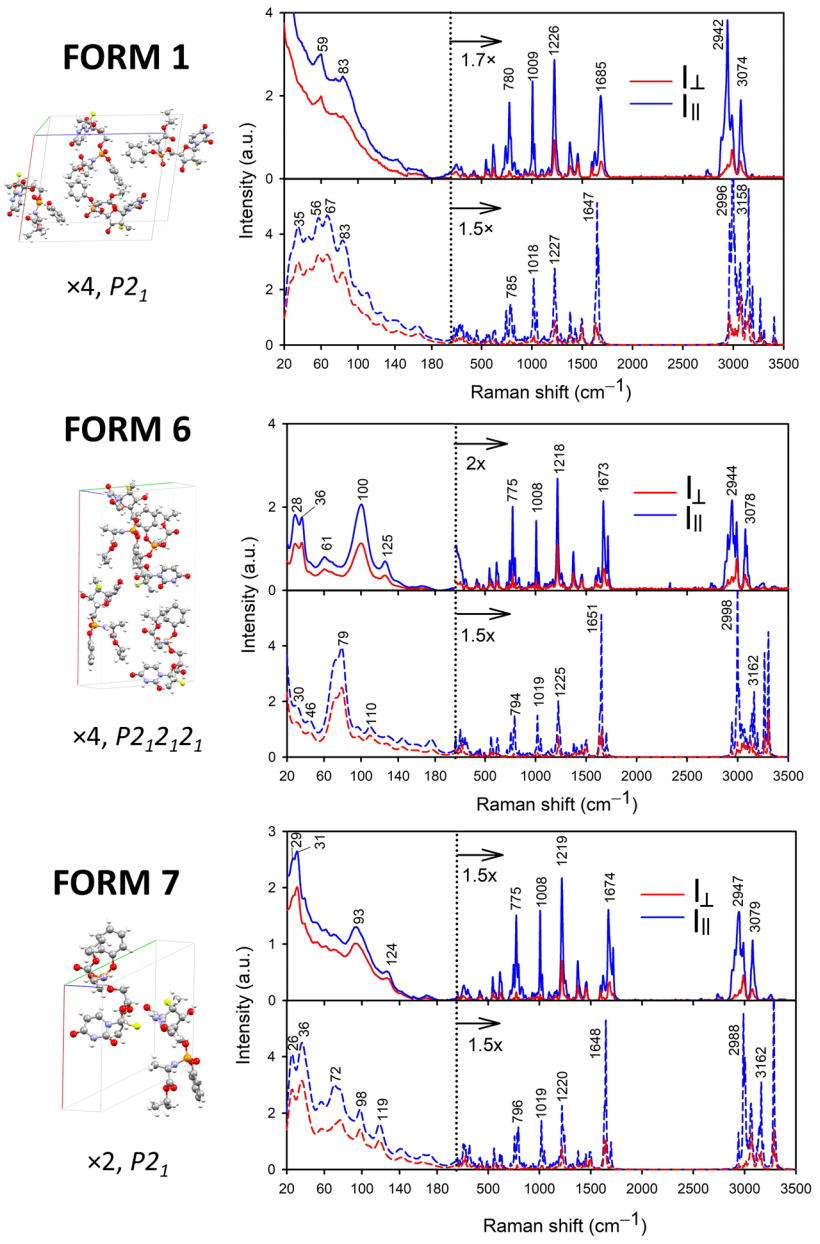


Figure 7. Experimental (upper) and calculated (lower) LPRS spectra of sofosbuvir forms 1, 6, and 7. Intensity is normalized to the ~775 cm^{-1} band.

The DOC ratios make the distinction possible even for forms 6 and 7, which exhibit almost identical unpolarized and similar linearly polarized Raman spectra. For instance, DOC(R) at 935 cm^{-1} is positive for each form but varies in intensity and shape. DOC(R) at 1125 cm^{-1} is even more specific, as it is positive for form 1 and negative for form 6 and form 7. Moreover, the 1125 cm^{-1} DOC of form 6 is approximately twice as intense as in the case of form 7. The DOC at ~1712 cm^{-1} has a broad profile for form 1 and is sharp for forms 6 and 7; form 6 additionally exhibits a shoulder close to the band at 1712 cm^{-1} at a higher wavenumber that is not present for forms 1 and 7. Similarly, the DOC profile varies for the individual forms around 3185 cm^{-1}. Thus, the 935/1125/1712/3185 cm^{-1} DOC markers appear quite helpful in differentiating between the polymorphs.

The raw I_L^L, I_R^L, $(I_L^L + I_R^L)$ and $(I_R^L - I_L^L)$ intensities are provided in Figure SV. I_R^R and I_L^R spectra are not included, the reason being that they are almost identical to the I_L^L and I_R^L counterparts. Calculated (PW, zero phonon mode) DOC spectra are presented in Figure SVI.

Effect of Scanned Area. To confirm that the Raman signal was averaged properly, we obtained the difference $(I_R^L - I_L^L)$ spectra of form 1 under different conditions. We scanned three regions of different size (300 × 300 μm, 100 × 100 μm and 10 × 10 μm) for two ratios of sofosbuvir to KBr (1:1 and 1:100). As seen in Figure SVII, the average $(I_R^L - I_L^L)$ spectrum acquired over a 100 × 100 μm scan area from pellets containing a 1:1 ratio of sofosbuvir to KBr matched that acquired across a 300 × 300 μm scan area. However, the spectrum obtained from 1:100 pellets over a 300 × 300 μm scan area had different relative intensities. Therefore, even for a large scanned area, the low API concentration did not provide reliable spectra. The spatial/area anisotropy was even more pronounced for form 6 (Figure SVIII), which inherently forms

https://doi.org/10.1021/acs.analchem.4c03573
Anal. Chem. 2024, 96, 18983−18993

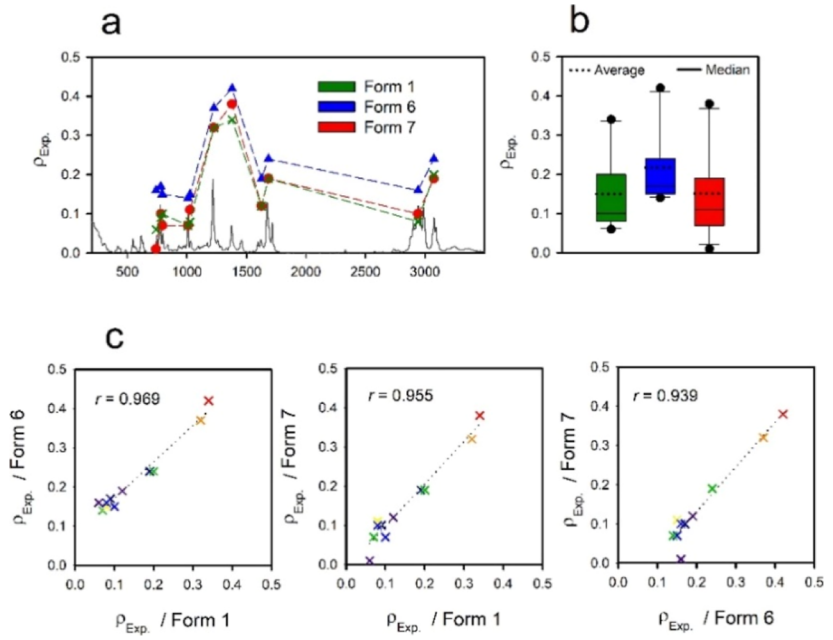


Figure 8. Experimental depolarization ratios ($\rho_{Exp.}$) of ten Raman bands with the most significant differences between the three forms of sofosbuvir [panel (a)]. Panel (b) summarizes the distributions of depolarization ratios, panel (c)-correlation coefficients across polymorph pairs.

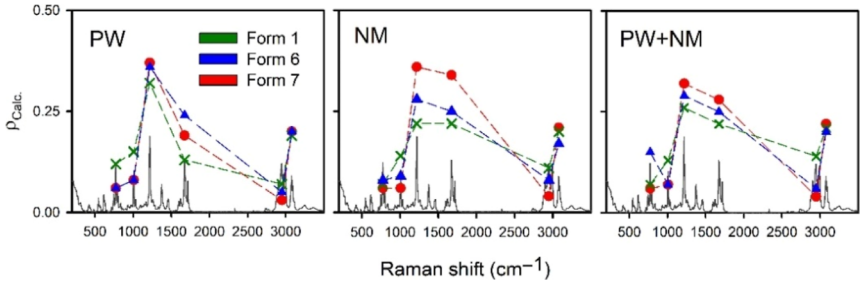


Figure 9. Calculated (0, 0, 0) phonon mode depolarization ratios ($\rho_{Calc.}$) of six Raman bands for sofosbuvir form 1, 6, and 7. The plotted Raman spectrum corresponds to form 6.

the largest crystals.[20] They are spread across the scanned area, and the $(I_R^L - I_L^L)$ spectrum from 1:100 pellets and a 300 × 300 μm scan area did not match the spectral pattern or relative intensities observed for more concentrated samples.

Spectral Similarity Indices of Experimental LPRS and CPRS Spectra. For an objective comparison, we calculated spectral similarity indices (σ)[70] in the 200−1800 cm^{-1} region (Figure 11) as

$$\sigma = \int S_{exp,1} S_{exp,2} d\nu / \left(\int S_{exp,1}^2 d\nu \int S_{exp,2}^2 d\nu \right)^{1/2}$$

For the same spectra, σ = 1, smaller values indicate disagreement. In LPRS, for $I_\perp$ σ vary by ∼2 to 6% among the three forms whereas for $I_\parallel$, σ vary by 1−6%. As expected, the highest values were obtained between the unpolarized spectra of forms 6 and 7, σ = 0.990. Comparison between form 6 and form 7 yields a value of 0.983. Using this indicator, in the fingerprint region, the $I_\perp$ spectra can differentiate between form 6 and form 7 about twice as efficiently as unpolarized spectra. CPRS offers even better distinction power than LPRS. The I_L^L spectra differ by ∼3 to 6% and the I_R^L spectra differ by ∼4 to 7%. For forms 6 and 7, $\sigma_{I_L^L} = 0.968$, which is three times greater compared to σ of unpolarized spectra. This makes the $(I_R^R - I_L^R)$, $(I_R^L - I_L^L)$ spectra DOCs convenient for polymorph discrimination, as the bisignate character of these spectra better highlights variations among chemical species. This could, for example, facilitate polymorph differentiation in industrial quality control.

The LPRS spectra in the LF region (20−200 cm^{-1}, Figure SIX) made the distinction even more reliable, as the unpolarized spectra between form 6 and form 7 give σ_{I_u} = 0.953 whereas the parallel ones give $\sigma_{I_\parallel}$ = 0.932. Nevertheless, the σs calculated for the stretching region (2700−3400 cm^{-1}, see Figure SX for LPRS and Figure SXI for CPRS) are similar to values for the fingerprint region.

CONCLUSION

To our knowledge, this study is the first to demonstrate the potential of LPRS and CPRS for differentiating drug polymorphs. We investigated the efficiency of unpolarized, linearly polarized and circularly polarized Raman spectra at resolving polycrystalline sofosbuvir forms 1, 6, and 7 in a broad spectral region (20−3500 cm^{-1}). By calculating the spectra at the DFT level, we were able to assign the main spectral features to their corresponding vibrational modes, including LF ones.

Although the unpolarized Raman spectra of forms 6 and 7 were nearly indistinguishable in the fingerprint and CH

https://doi.org/10.1021/acs.analchem.4c03573
Anal. Chem. 2024, 96, 18983−18993

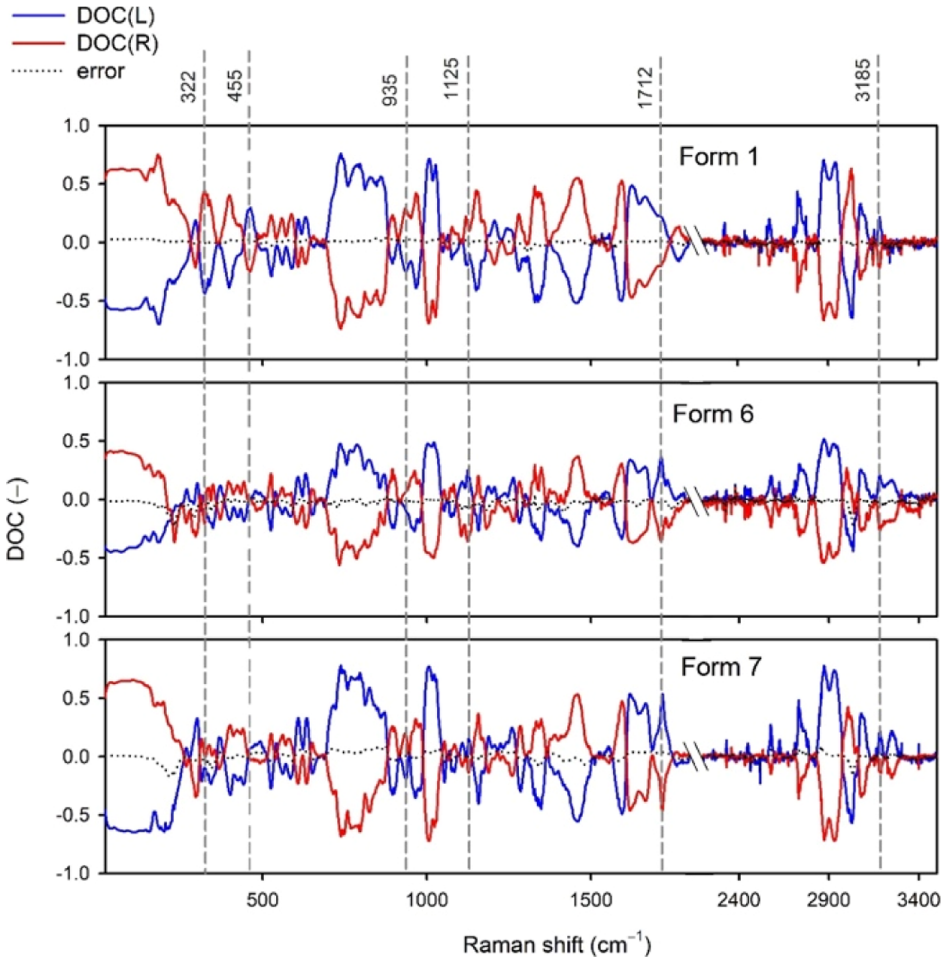


Figure 10. Experimental DOC ratios of sofosbuvir forms. The most prominent differences are marked with vertical dashed lines. The region between 1800 and 2200 cm^{-1} is omitted.

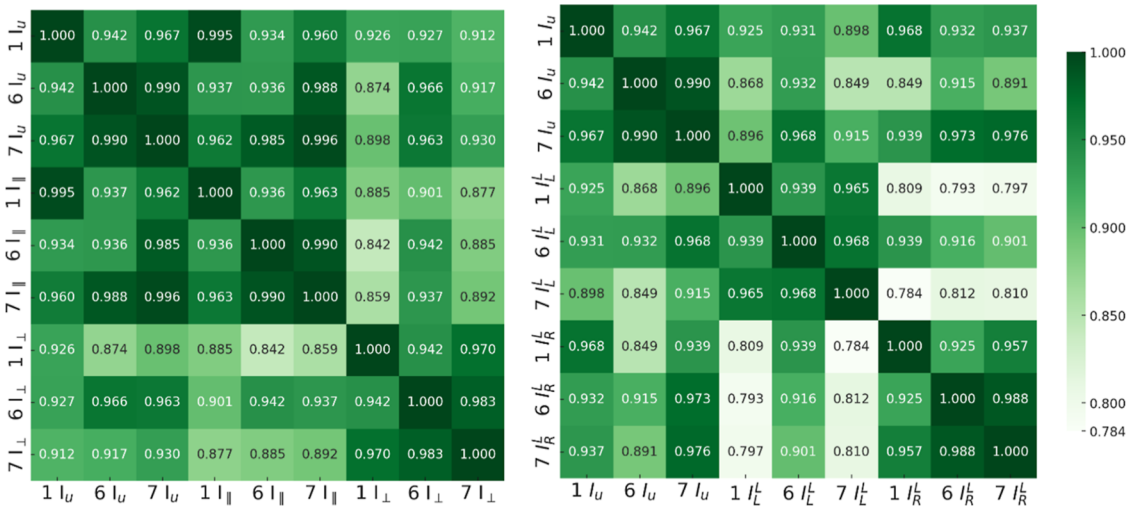


Figure 11. "Heat maps" of spectral similarity indices between experimental LPRS (left) and CPRS (right) spectra. The indices are calculated for integrals within the range of 200−1800 cm^{-1}.

stretching regions, the LF one provided spectral patterns distinct for each form. We observed that the intensities of LF bands were approximately three times greater than for typical fingerprint bands. Based on calculations, we attributed a large part of the intensity enhancement to the vibrations of highly polarizable phenyl and uracil rings.

In contrast to unpolarized Raman spectroscopy, its linearly polarized variant was able to distinguish between forms 6 and 7 even in the fingerprint region and the CH-stretching region. Form 6 exhibited the highest values of ρ_{Exp}, forms 1 and 7 gave similar $\rho_{Exp.}$ values. CPRS further enhanced the distinction. DOC ratios at 935, 1125, 1712, and 3185 cm^{-1} were proposed

https://doi.org/10.1021/acs.analchem.4c03573
Anal. Chem. 2024, 96, 18983−18993

as the best markers for distinguishing polymorphs of sofosbuvir.

By calculating spectral similarity indices, we could better quantify the performance of unpolarized, LPRS and CPRS spectra in resolving sofosbuvir polymorphs. For example, in the fingerprint region, LPRS was approximately twice as effective and CPRS three times as effective as the unpolarized spectra at differentiating the structurally similar forms 6 and 7. Both types of polarized Raman spectroscopy can be effectively used in simplified or automated quality control of pharmaceuticals.

■ ASSOCIATED CONTENT

Data Availability Statement

The data sets generated during the current study are available via the OSF data repository (doi: 10.17605/OSF.IO/S45DU; https://osf.io/s45du).

Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acs.analchem.4c03573.

Additional details of the procedures can be found (PDF)

■ AUTHOR INFORMATION

Corresponding Author

Jakub Kaminský − *Institute of Organic Chemistry and Biochemistry of the Academy of Sciences, Prague 16610, Czech Republic;* orcid.org/0000-0001-6347-3022; Email: jakub.kaminsky@uochb.cas.cz

Authors

Věra Schrenková − *Institute of Organic Chemistry and Biochemistry of the Academy of Sciences, Prague 16610, Czech Republic; University of Chemistry and Technology Prague, Prague 16628, Czech Republic;* orcid.org/0000-0002-8882-9627

Josef Kapitán − *Palacký University Olomouc, Olomouc 77146, Czech Republic;* orcid.org/0000-0002-1916-9186

Petr Bouř − *Institute of Organic Chemistry and Biochemistry of the Academy of Sciences, Prague 16610, Czech Republic;* orcid.org/0000-0001-8469-1686

Argyro Chatziadi − *University of Chemistry and Technology Prague, Prague 16628, Czech Republic;* orcid.org/0000-0002-4004-6818

Adam Sklenář − *Institute of Organic Chemistry and Biochemistry of the Academy of Sciences, Prague 16610, Czech Republic; University of Chemistry and Technology Prague, Prague 16628, Czech Republic*

Complete contact information is available at: https://pubs.acs.org/10.1021/acs.analchem.4c03573

Author Contributions

V.S.—conceptualization, Raman experiments, calculations, writing; J.Kap.—Raman experiment supervision; P.B.—supervision, programming, writing; A.Ch.—preparation of polymorphs; A.S.—calculations, sample preparation; J.Kam.—supervision, calculations, writing.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The work was supported by the Czech Science Foundation (24-10558S) and the University of Chemistry and Technology, Prague (VIGA grant A2_FCHI_2024_028). Computational resources were provided by the e-INFRA CZ project (ID: 90254) supported by the Ministry of Education, Youth and Sports of the Czech Republic.

■ REFERENCES

(1) Fontana, F.; Figueiredo, P.; Zhang, P.; Hirvonen, J. T.; Liu, D.; Santos, H. A. *Adv. Drug Delivery Rev.* **2018**, *131*, 3−21.

(2) Štejfa, V.; Bazyleva, A.; Fulem, M.; Rohlíček, J.; Skořepová, E.; Růžička, K.; Blokhin, A. V. *J. Chem. Thermodyn.* **2019**, *131*, 524−543.

(3) Bharate, S. S. *Drug Discovery Today* **2021**, *26*, 384−398.

(4) Zhu, B.; Zhang, Q.; Ren, G.; Mei, X. *Cryst. Growth Des.* **2017**, *17*, 5994−6005.

(5) Zhou, Y.; Lv, C.; Liu, X.; Gao, J.; Gao, Y.; Wang, T.; Huang, X. *Cryst. Growth Des.* **2024**, *24*, 584−600.

(6) Tandon, R.; Tandon, N.; Thapar, R. K. *Pharm. Pat. Anal.* **2018**, *7*, 59−63.

(7) Raw, A. S.; Furness, M. S.; Gill, D. S.; Adams, R. C.; Holcombe, F. O.; Yu, L. X. *Adv. Drug Delivery Rev.* **2004**, *56*, 397−414.

(8) Censi, R.; Di Martino, P. *Molecules* **2015**, *20*, 18759−18776.

(9) Sládková, V.; Dammer, O.; Sedmak, G.; Skořepová, E.; Kratochvíl, B. *Crystals* **2017**, *7*, 13.

(10) Rodríguez, I.; Gautam, R.; Tinoco, A. D. *Biomimetics* **2021**, *6*, 1.

(11) Du, Y.; Su, Y. *Solid State Nucl. Magn. Reson.* **2022**, *120*, 101796.

(12) Mathew, R.; Uchman, K. A.; Gkoura, L.; Pickard, C. J.; Baias, M. *Magn. Reson. Chem.* **2020**, *58*, 1018−1025.

(13) Wang, L.-L.; Wang, L.-Y.; Yu, Y.-M.; Li, Y.-T.; Wu, Z.-Y.; Yan, C.-W. *CrystEngComm* **2020**, *22*, 5010−5021.

(14) Shah, K. C.; Shah, M. B.; Solanki, S. J.; Makwana, V. D.; Sureja, D. K.; Gajjar, A. K.; Bodiwala, K. B.; Dhameliya, T. M. *J. Mol. Struct.* **2023**, *1278*, 134914.

(15) Inoue, M.; Hisada, H.; Koide, T.; Fukami, T.; Roy, A.; Carriere, J.; Heyler, R. *Anal. Chem.* **2019**, *91*, 1997−2003.

(16) Park, H.; Nie, H.; Dhiman, A.; Tomar, V.; Zhou, Q. T. *Mol. Pharm.* **2020**, *17*, 3043−3052.

(17) Dan Córdoba, A. V.; Aiassa, V.; Dimmer, J. A.; Barrionuevo, C. N.; Quevedo, M. A.; Longhi, M. R.; Zoppi, A. *Pharmaceutics* **2023**, *15*, 198.

(18) Frelek, J.; Górecki, M.; Łaszcz, M.; Suszczyńska, A.; Vass, E.; Szczepek, W. J. *Chem. Commun.* **2012**, *48*, 5295−5297.

(19) Frelek, J.; Gorecki, M.; Dziedzic, A.; Jablonska, E.; Kamienski, B.; Wojcieszczyk, R. K.; Luboradzki, R.; Szczepek, W. J. *J. Pharm. Sci.* **2015**, *104*, 1650−1657.

(20) Sklenář, A.; Růžičková, L.; Schrenková, V.; Bednárová, L.; Pazderková, M.; Chatziadi, A.; Zmeškalová Skořepová, E.; Šoóš, M.; Kaminský, J. *Spectrochim. Acta, Part A* **2024**, *318*, 124478.

(21) Haefele, T. F.; Paulus, K. Confocal raman microscopy in pharmaceutical development. In *Springer Series in Optical Sciences*; Dieing, T., Hollricher, O., Toporski, J., Eds.; Springer, 2010; Vol. *158*, pp 165−202.

(22) Luebbert, C.; Klanke, C.; Sadowski, G. *Int. J. Pharm.* **2018**, *535*, 245−252.

(23) Reddy, J. P.; Jones, J. W.; Wray, P. S.; Dennis, A. B.; Brown, J.; Timmins, P. *Int. J. Pharm.* **2018**, *541*, 253−260.

(24) Carriere, J. T.; Havermeyer, F. Ultra-low frequency Stokes and anti-Stokes Raman spectroscopy at 785nm with volume holographic grating filters. *Progress in Biomedical Optics and Imaging-Proceedings of SPIE*, 2012; Vol. *8219*.

(25) Kim, J.; Lee, J. U.; Cheong, H. *J. Phys.: Condens. Matter* **2020**, *32*, 343001.

(26) Bērziņš, K.; Fraser-Miller, S. J.; Gordon, K. C. *Int. J. Pharm.* **2021**, *592*, 120034.

(27) Hatipoglu, M. K.; Zaker, Y.; Willett, D. R.; Gupta, N.; Rodriguez, J. D.; Patankar, S.; Capella, P.; Yilmaz, H. *Anal. Chem.* **2023**, *95*, 15325−15332.

(28) Ruggiero, M. T.; Sutton, J. J.; Fraser-Miller, S. J.; Zaczek, A. J.; Korter, T. M.; Gordon, K. C.; Zeitler, J. A. *Cryst. Growth Des.* **2018**, *18*, 6513−6520.

(29) Aviv, H.; Nemtsov, I.; Mastai, Y.; Tischler, Y. R. *J. Phys. Chem. A* **2017**, *121*, 7882−7888.
(30) Ge, M.; Wang, Y.; Zhu, J.; Wu, B.; Xu, D.; Yao, J. *Sensors* **2022**, *22*, 4053.
(31) Bērziņš, K.; Sales, R. E.; Barnsley, J. E.; Walker, G.; Fraser-Miller, S. J.; Gordon, K. C. *Vibr. Spectrosc.* **2020**, *107*, 103021.
(32) Xu, B.; Mao, N.; Zhao, Y.; Tong, L.; Zhang, J. *J. Phys. Chem. Lett.* **2021**, *12*, 7442−7452.
(33) Thyr, J.; Österlund, L.; Edvinsson, T. *J. Raman Spectrosc.* **2021**, *52*, 1395−1405.
(34) Zou, B.; Wei, Y.; Zhou, Y.; Ke, D.; Zhang, X.; Zhang, M.; Yip, C.-T.; Chen, X.; Li, W.; Sun, H. *Nanoscale Horiz.* **2021**, *6*, 809−818.
(35) Ilchenko, O.; Pilgun, Y.; Kutsyk, A.; Bachmann, F.; Slipets, R.; Todeschini, M.; Okeyo, P. O.; Poulsen, H. F.; Boisen, A. *Nat. Commun.* **2019**, *10*, 5555.
(36) Motai, K.; Koishihara, N.; Narimatsu, T.; Ohtsu, H.; Kawano, M.; Wada, Y.; Akisawa, K.; Okuwaki, K.; Mori, T.; Kim, J.-S.; et al. *Cryst. Growth Des.* **2023**, *23*, 4556−4561.
(37) Zhang, S.; Zhang, N.; Zhao, Y.; Cheng, T.; Li, X.; Feng, R.; Xu, H.; Liu, Z.; Zhang, J.; Tong, L. *Chem. Soc. Rev.* **2018**, *47*, 3217−3240.
(38) Fasolato, C.; Zardo, I.; De Luca, M. Addressing Crystal Structure in Semiconductor Nanowires by Polarized Raman Spectroscopy. In *Fundamental Properties of Semiconductor Nanowires*; Fukata, N., Rurali, R., Eds.; Springer: Singapore, 2021; pp 307−348.
(39) Moriyama, K.; Onishi, H.; Ota, H. *Pharm. Anal. Acta* **2015**, *6*, 1−3.
(40) Kolesov, B. A.; Mikhailenko, M. A.; Boldyreva, E. V. *Phys. Chem. Chem. Phys.* **2011**, *13*, 14243−14253.
(41) Zhao, Y.; Han, S.; Zhang, J.; Tong, L. *J. Raman Spectrosc.* **2021**, *52*, 525−531.
(42) Saito, R.; Hung, N. T.; Yang, T.; Huang, J.; Liu, H.-L.; Gulo, D. P.; Han, S.; Tong, L. *Small* **2024**, 2308558.
(43) Han, S.; Hung, N. T.; Xie, Y.; Saito, R.; Zhang, J.; Tong, L. *Nano Lett.* **2023**, *23*, 8454−8459.
(44) Hug, W.; Hangartner, G. *J. Raman Spectrosc.* **1999**, *30*, 841−852.
(45) Guo, S.; Popp, J.; Bocklitz, T. *Nat. Protoc.* **2021**, *16*, 5426−5459.
(46) Milani, A. *J. Phys. Chem. B* **2015**, *119*, 3868−3874.
(47) Pagliai, M.; Osticioli, I.; Nevin, A.; Siano, S.; Cardini, G.; Schettino, V. *J. Raman Spectrosc.* **2018**, *49*, 668−683.
(48) Schrenková, V.; Para Kkadan, M. S.; Kessler, J.; Kapitán, J.; Bouř, P. *Phys. Chem. Chem. Phys.* **2023**, *25*, 8198−8208.
(49) Bērziņš, K.; Sutton, J. J.; Fraser-Miller, S. J.; Rades, T.; Korter, T. M.; Gordon, K. C. *Cryst. Growth Des.* **2020**, *20*, 6947−6955.
(50) Davis, M. P.; Korter, T. M. *Mol. Pharm.* **2022**, *19*, 3385−3393.
(51) Clark, S. J.; Segall, M. D.; Pickard, C. J.; Hasnip, P. J.; Probert, M. I. J.; Refson, K.; Payne, M. C. *Z. für Kristallogr. -Cryst. Mater.* **2005**, *220*, 567−570.
(52) Bouř, P.; Sopková, J.; Bednárová, L.; Maloň, P.; Keiderling, T. A. *J. Comput. Chem.* **1997**, *18*, 646−659.
(53) Hoshina, H.; Ishii, S.; Yamamoto, S.; Morisawa, Y.; Sato, H.; Uchiyama, T.; Ozaki, Y.; Otani, C. *IEEE Trans. Terahertz Sci. Technol.* **2013**, *3*, 248−258.
(54) Dračínský, M.; Procházková, E.; Kessler, J.; Šebestík, J.; Matějka, P.; Bouř, P. *J. Phys. Chem. B* **2013**, *117*, 7297−7307.
(55) Nakamura, M.; Kanda, T.; Haga, Y.; Sasaki, R.; Wu, S.; Nakamoto, S.; Yasui, S.; Arai, M.; Imazeki, F.; Yokosuka, O. *World J. Hepatol.* **2016**, *8*, 183−190.
(56) Mande, H. M.; Jetti, R. K. R.; Das, P. P.; Nicolopoulos, S. *Acta Crystallogr., Sect. A: Found. Adv.* **2017**, *73*, C434.
(57) Chatziadi, A.; Skorepova, E.; Jirat, J.; Rohlíček, J.; Šoóš, M. *Cryst. Growth Des.* **2022**, *22* (5), 3395−3404.
(58) Chatziadi, A.; Skořepová, E.; Rohlíček, J.; Dušek, M.; Ridvan, L.; Šoóš, M. *Cryst. Growth Des.* **2020**, *20*, 139−147.
(59) Program sc95. *Institute of Organic Chemistry and Biochemistry of the CAS*: Prague, Czech Republic, 2024.
(60) Krupová, M.; Kessler, J.; Bouř, P. *ChemPhysChem* **2021**, *22*, 83−91.
(61) Program QGRAD. *Institute of Organic Chemistry and Biochemistry of the CAS*: Prague, Czech Republic, 2020.
(62) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Petersson, G. A.; Nakatsuji, H.; Li, X.; Caricato, M.; Marenich, A. V.; Bloino, J.; Janesko, B. G.; Gomperts, R.; Mennucci, B.; Hratchian, H. P.; Ortiz, J. V.; Izmaylov, A. F.; Sonnenberg, J. L.; Williams-Young, D.; Ding, F.; Lipparini, F.; Egidi, F.; Goings, J.; Peng, B.; Petrone, A.; Henderson, T.; Ranasinghe, D.; Zakrzewski, V. G.; Gao, J.; Rega, N.; Zheng, G.; Liang, W.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Throssell, K.; Montgomery, J. A.; Peralta, J. E.; Ogliaro, F.; Bearpark, M. J.; Heyd, J. J.; Brothers, E. N.; Kudin, K. N.; Staroverov, V. N.; Keith, T. A.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A. P.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Millam, J. M.; Klene, M.; Adamo, C.; Cammi, R.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Farkas, O.; Foresman, J. B.; Fox, D. J. *Gaussian 16*, Revision C.01; Gaussian, Inc.: Wallingford CT, 2016.
(63) Bouř, P.; Keiderling, T. A. *J. Chem. Phys.* **2002**, *117*, 4126−4132.
(64) Takano, Y.; Houk, K. N. *J. Chem. Theory Comput.* **2005**, *1*, 70−77.
(65) Daněček, P.; Bouř, P. *J. Comput. Chem.* **2007**, *28*, 1617−1624.
(66) Roy, S.; Chamberlin, B.; Matzger, A. J. *Org. Process Res. Dev.* **2013**, *17*, 976−980.
(67) Ayala, A. P. *Vibr. Spectrosc.* **2007**, *45*, 112−116.
(68) Larkin, P. J.; Dabros, M.; Sarsfield, B.; Chan, E.; Carriere, J. T.; Smith, B. C. *Appl. Spectrosc.* **2014**, *68*, 758−776.
(69) Ohno, K.; Kimura, J.; Yamakita, Y. *Chem. Phys. Lett.* **2001**, *342*, 207−219.
(70) Debie, E.; De Gussem, E.; Dukor, R. K.; Herrebout, W.; Nafie, L. A.; Bultinck, P. *ChemPhysChem* **2011**, *12*, 1542−1549.

https://doi.org/10.1021/acs.analchem.4c03573
Anal. Chem. 2024, 96, 18983−18993