Journal of Molecular Modeling (2022) 28:350
https://doi.org/10.1007/s00894-022-05340-3

ORIGINAL PAPER

Fluorinated derivatives of tetrahydroaltersolanol molecule on COVID-19, HIV, and HTLV protease by DFT and molecular docking approaches

Maliheh Azadparvar[1] · M. Kheirabadi[2] · H. A. Rahnamaye Aliabad[1]

Received: 26 April 2022 / Accepted: 28 September 2022

Abstract
Structural, optoelectronic, and biological properties of tetrahydroaltersolanol ($C_{16}H_{20}O_7$) and fluorinated derivatives are calculated using density functional theory (DFT) and molecular docking approaches. It is shown that the pure $C_{16}H_{20}O_7$ molecule has a direct HOMO–LUMO energy gap about 3.1 eV. The substitution of F atom at A category decreases the electronic energy gap, while it is constant at B category. In A category, the behavior of the pure molecule changed from insulator to semiconductor with various substitution of F atom. The electronic properties were depended on the F sites in the pure molecule. The molecular electrical transport properties and charge-transfer possibilities increase with decreasing energy gap. The pure $C_{16}H_{20}O_7$ molecule with high energy gap has low chemical reactivity and substitution of F atom at considered molecule increases chemical reactivity. Obtained results show that F-O bonds in trifurcation bonds of $C_{16}H_{19}O_7$(F14), $C_{16}H_{19}O_7$(F16), and $C_{16}H_{19}O_7$(F17) molecules play a key role in confronting with COVID-19, HIV, and HTLV proteases, respectively. Optical spectra, such as the dielectric functions, electron energy-loss spectroscopy, refractive index, extinction coefficient, and reflection spectra show that fluorinated derivatives of $C_{16}H_{20}O_7$ at B category can be used in the new drugs.

Keywords DFT · Molecular docking · Fluorinated derivatives of tetrahydroaltersolanol

Introduction

Fungi are found in all environments with a universal spread and grow in different habitats [1–3]. Hydroanthraquinones, alkaloids, polyketides, and peptides are bioactive materials which produce by species of fungi and some of them are toxic to plants, insects and animals. Fungal metabolites have various applications in medicine as antiviral, antibacterial, and anticancer compounds [4–6]. The reduced species of anthraquinones that are formed in fungi such as *Alternaria solani* [7, 8], *A. porri* [9], *Dactylaria lutea* [10], *Phomopsis juniperivora* [11], and *Stemphylium botryosum* [12] are called altersolanol, which contains many different types and exhibits interesting properties. For instance, altersolanols exhibited significant antimicrobial, antiprotozoal, and cytotoxic activities [13]. Marine fungi have recently emerged as a newfound source of fungal natural products and maybe have a potentially crucial point in drug discovery.

Genetic material (DNA and RNA) of the virus encodes viral protease as essential enzymes to complete the viral infections cycle. These enzymes cleave selectively the scissile peptide bonds with different catalytic mechanisms in viral or cellular protein. Serine, cysteine, or aspartic acid residues in the active site of viral proteases recognize the specific peptide bond that is often located within conserved sequence motifs. The 2019-nCOV as novel coronavirus has main cysteine protease named (M^{pro}, 3C) [14], HIV and HTLV protease [15]. The aspartic acid protease family have potential drug targets because of their essential activities. One of the main methods commonly used to study the physical and chemical properties of materials is density functional theory (DFT) approach [16] as implemented in the Wien2k code [17]. Fluorine with high electronegativity in carbon–fluorine bond plays a key role to create a

Maliheh Azadparvar and M. Kheirabadi contributed equally to this work.

✉ H. A. Rahnamaye Aliabad
Rahnamaye@gmail.com; Rahnama@hsu.ac.ir

[1] Department of Physics, Hakim Sabzevari University, Sabzevar 96179-76487, Iran

[2] Department of Biology, Hakim Sabzevari University, Sabzevar 96179-76487, Iran

Published online: 11 October 2022

strongly polarized bond with the highest energies. A hydrogen atom or a hydroxyl group can displace with fluorine [18, 19]. Fluorine is rarely found in natural compounds, and monofluorinated compounds are produced by some plants and bacteria [20, 21]. They often play a key role in defense strategy due to their toxic properties. Nowadays, fluorinated compounds have various applications in agriculture and medicine [22]. On the other hand, hydroanthraquinone derivatives as bioactive metabolites were obtained from a soft coral–derived *Alternaria* sp. fungus [23]. One of them is a tetrahydroaltersolanol with $C_{16}H_{20}O_7$ formula and SMILES (O = C1c2c(O)cc(cc2[C@@H](O)[C@@H]2[C@H] (O) [C@@] (O)([C@@H] (O)C[C@@H]12) C) OC) exhibited antiviral activity against porcine reproductive and respiratory syndrome virus (PRRSV) with IC_{50} values 65 $\mathscr{M}$M [23].

In this study, we have investigated the structural, optoelectronic, and biological properties of tetrahydroaltersolanol ($C_{16}H_{20}O_7$) [23] and fluorinated derivatives by using density functional theory (DFT) and molecular docking approaches. The 19 fluorinated derivatives of tetrahydroaltersolanol are constructed via hydrogen substitution by fluorine. The optoelectronic and ADME properties were predicted by the Wien2k code and Molsoft web server, respectively. Finally, binding mode of tetrahydroaltersolanol and designed compounds to HIV, HTLV protease, and COVID-19 protease were obtained by molecular docking and compared with the obtained DFT results.

Computational details

A. The FP-LAPW method as implemented in WIEN2k

In the first principal calculations, the full-potential linearized augmented plane wave method (FP- LAPW) is used for simulation of the structural and optoelectronic properties of pure $C_{16}H_{20}O_7$ and $C_{16}H_{19}O_7$(F1) to $C_{16}H_{19}O_7$(F19) compounds by Wien2k code [16, 17, 24]. The exchange–correlation potentials are calculated using the generalized gradient approximation (GGA). The tetrahydroaltersolanol ($C_{16}H_{20}O_7$) compound is crystalized at orthorhombic structure with the "P 21 21 21" space group, as shown in Fig. 1. The lattice constants are a = 5.6702 Å, b = 10.8579 Å, and c = 23.9550 Å. The number of k-points is selected about 1000 for self-consistent calculations. In the FP- LAPW calculations, the cut-off energy, which defines the separation of core and valence states is also selected − 6 Ry with the convergence parameter of 0.0001 Ry.

B. Molecular docking and molecular dynamic

Molecular docking is a computation process by which ligands are moved in a 3D space so that, finally, the best space structures of the target and ligand positions are found by getting the highest scores in the functions. Before docking process, to predict drug-likeness of compounds, ADME properties of the compounds were predicted using Molsoft web services (https://molsoft.com/mprop/) which estimates the absorption, distribution, metabolism, and excretion of candidate compounds. For docking process, proteins and ligands optimized. At the first step, primary ligand optimization was performed by Hyperchem software [25], a salvation box was created for them, and molecular dynamic simulation was done by amber force field [25] under the temperature of 300 K in 10 picoseconds in order to achieve the most sustainable state and the lowest energy of the structure of inhibitors for being used in the future studies. Then, these structures were saved as PDB files to be used in the docking process [26]. At the second step, the crystallography structure of HIV, HTLV, and COVID-19 protease enzymes were selected from the protein data bank with the respective PDB codes of 1YT9, 3LIN, and 6M03. Then, all the non-amino acid molecules (water, ion, ligand, inhibitor, etc.)

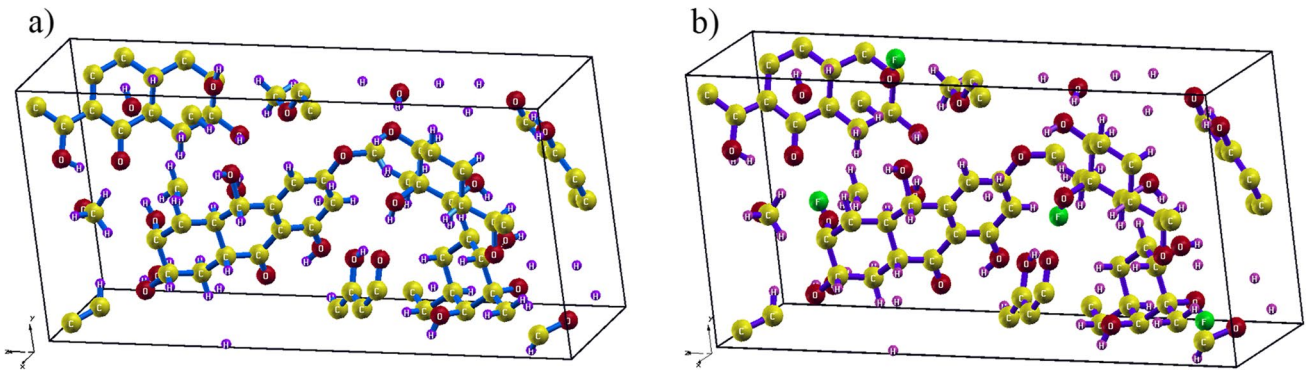

Fig. 1 The unit cells of **a** $C_{16}H_{20}O_7$ and **b** $C_{16}H_{19}O_7$(F1)

were removed. Then, the structure of the HTLV and HIV protease's optimization was done by Gromacs 5.1.2 software [27] in a quasi-natural condition in terms of temperature, atmosphere pressure, pH, and the existence of water as a solvent to get their balanced structure. First, the topology of protein was built by using the GROMOS96 53a5 force field. Then, a cubic box was considered between the protein and the edge of the box with the distance of 1.0 nm, and the water molecules were added to the box as solvent by using simple point charge (SPC) model. Then, four chlorine ions were added for neutralizing the system load. In the following, energy minimization was done by the steepest descent algorithm. Then, NVT equilibration and NPT equilibration were done each for 100 picoseconds, and finally, molecular dynamic simulation was done for 10,000 picoseconds (10 ns). The docking process was done by Auto Dock 4.2 software. In the docking process, LGA (Lamarckian genetic algorithm) was used and all the non-rotatable amide bonds were selected. The cubic box with the size of $126 \times 126 \times 126$ Å was set to cover all the protein–ligand complexes. The number of GA runs for each 250 docks was selected by 25,000,000 energy measurement [28]. The content of hydrogen bonds and hydrophobic interactions between ligand and protein in the complex was analyzed by PLIP server. Data were drawn with PyMol software [29].

Results and discussion

Today, in medicine, plants and fungi are used as rich sources of various metabolites to treat various diseases. The resulting metabolites can have inhibitory effects on the cancer cells or viral proteins. On the other hand, rare natural fluorinated compounds are produced by plants and bacteria that play a major role in defense mechanisms. These compounds have attracted special attention as effective drug compounds. Fluorinated derivatives of chemical compounds are usually toxic; still it should also be borne in mind that the enzymes that break down fluorinated compounds are found only in small numbers in some organisms. Therefore, the necessary precautions should be taken in their therapeutic uses. One of the hydroanthraquinone derivatives extracted from marine fungi with antiviral activity against porcine reproductive and respiratory syndrome virus (PRRSV) with IC_{50} values 65 micromolar (μM) [23] was selected to design its fluorinated derivatives. The hydrogen atoms in hydroanthraquinones (Cif code: 1,503,469) were substituted with fluorine (Fig. 2). As the first step to design new compounds in lab, they should be energetically stable. To estimate stability of designed fluorinated substance, electronic and optical properties of them was calculated. After that, drug likeness and ADME properties of fluorinated substances was predicted. Finally, binding mode of compounds to HTLV, HIV, and COVID-19 protease was obtained.

A. Optoelectronic results

The physical and chemical properties of materials are studied by investigating the behavior of electrons near the fermi level. The electrons in fermi level play a key role in the reactive properties of drugs. In this part, an attempt has been made to obtained correlation between DFT calculations and pharmaceutical properties. It is shown that several computational parameters such as highest occupied molecular orbital (HOMO) and lowest unoccupied molecular orbital (LUMO) band gap and absorption spectra were predicted in order to describe the pharmaceutical properties of tetrahydroaltersolanol ($C_{16}H_{20}O_7$) and fluorinated derivatives. The calculated

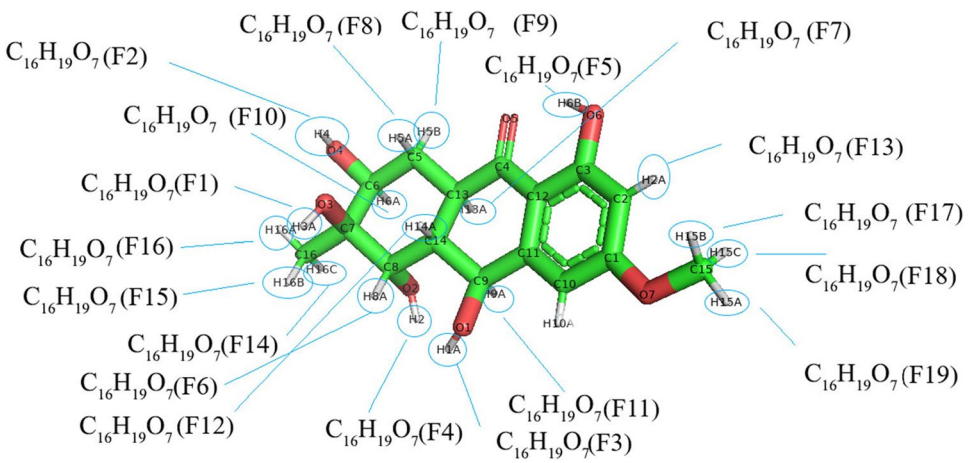

Fig. 2 Fluorinated derivatives of hydroanthraquinones. The figures were generated by PyMol [31]

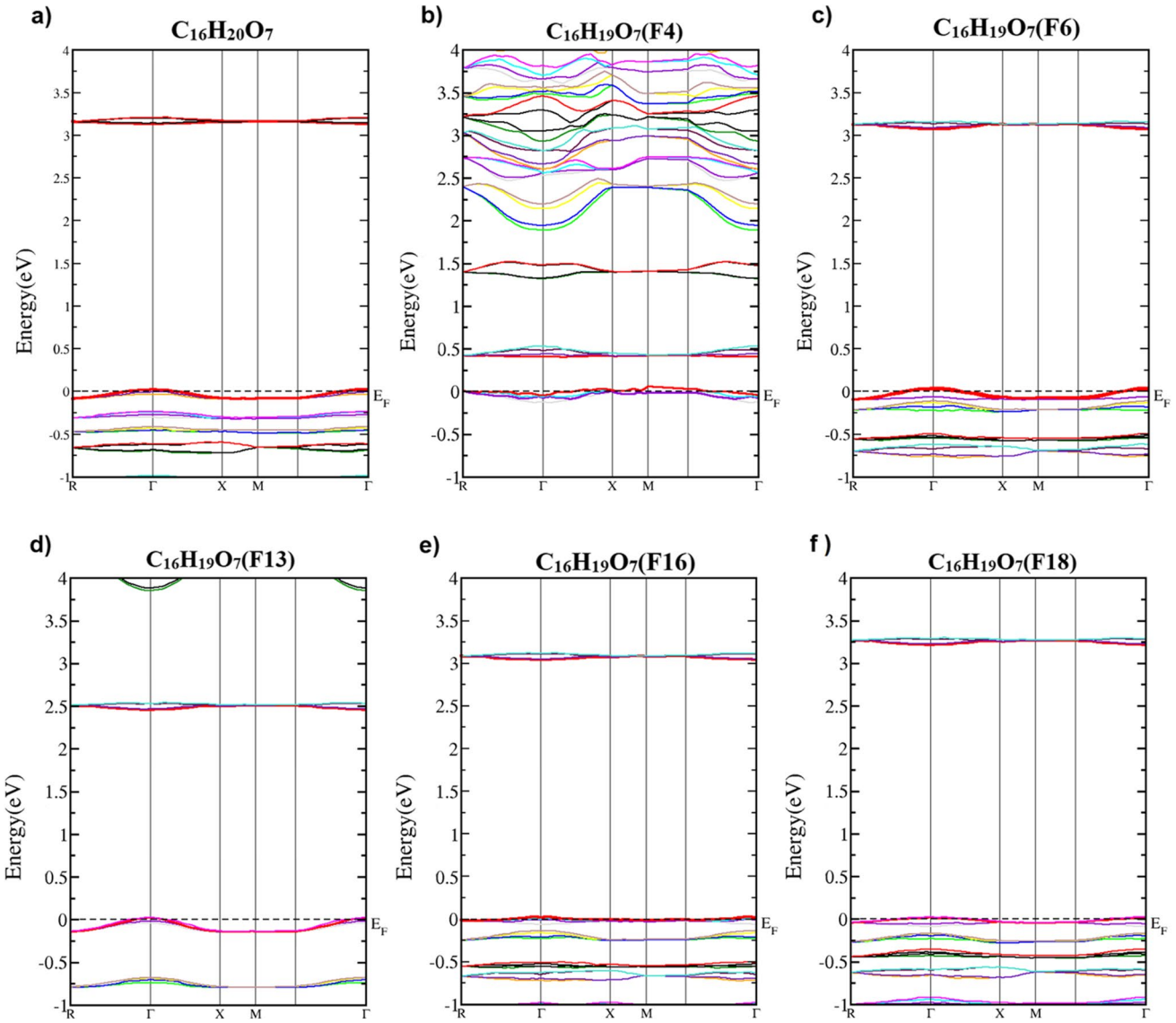

Fig. 3 The band structures of **a** $C_{16}H_{20}O_7$, **b** $C_{16}H_{19}O_7$(F4) from category A, **c** $C_{16}H_{19}O_7$(F6), **d** $C_{16}H_{19}O_7$(F13), **e** $C_{16}H_{19}O_7$(F16), and **f** $C_{16}H_{19}O_7$(F18) from category B. The black dash line shows Fermi energy level

electronic band structures and density of states of the $C_{16}H_{20}O_7$ change by the replacement of a hydrogen atom by a fluorine in $C_{16}H_{19}O_7$(F1 to F19) compounds, as shown in Figs. 3, 4, 5, and 6. The calculated bandgap values of all compounds are given in Table 1. For better presentation of the obtained results, all compounds are classified into one of two groups: category A or category B. Category A involves the fluorine atoms that joined to the oxygen atom, while in category B, the fluorine atom is attached to carbon that are located outside benzene rings with three hydrogen bonds or on benzene rings. The results show that the $C_{16}H_{20}O_7$ compound has a direct bandgap about 3.1 eV (Fig. 3). The $C_{16}H_{20}O_7$ bandgap is reduced upon the replacement of a hydrogen atom by a fluorine in the $C_{16}H_{19}O_7$ (F1 to F5) compounds. On the other hand, obtained bandgaps for $C_{16}H_{19}O_7$(F6 to F19) compounds, except for $C_{16}H_{19}O_7$(F13) compound, are direct and near to the bandgap value of pure one. By decreasing the electronic band gap, the behavior of the $C_{16}H_{20}O_7$ compound is altered from semiconductor to metal.

The HOMO–LUMO band gap describes the charge transfer interaction within the molecule. The pure tetrahydroaltersolanol ($C_{16}H_{20}O_7$) and the B category of fluorinated derivatives with a high HOMO–LUMO band gap values have low chemical reactivity and high kinetic

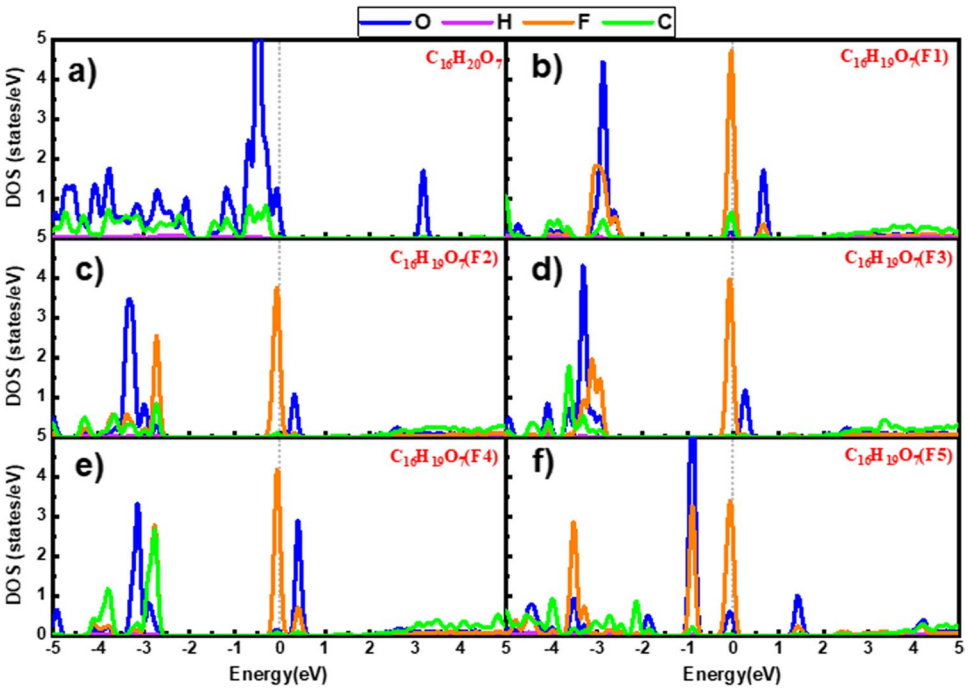

Fig. 4 The calculated total density of states for **a** $C_{16}H_{20}O_7$ and **b** $C_{16}H_{19}O_7$(F1), **c** $C_{16}H_{19}O_7$(F2), **d** $C_{16}H_{19}O_7$(F3), **e** $C_{16}H_{19}O_7$(F4), and **f** $C_{16}H_{19}O_7$(F5) at A category of compounds. In these compounds, the fluorine atom is bonded to oxygen atom. The gray dot line shows Fermi energy level

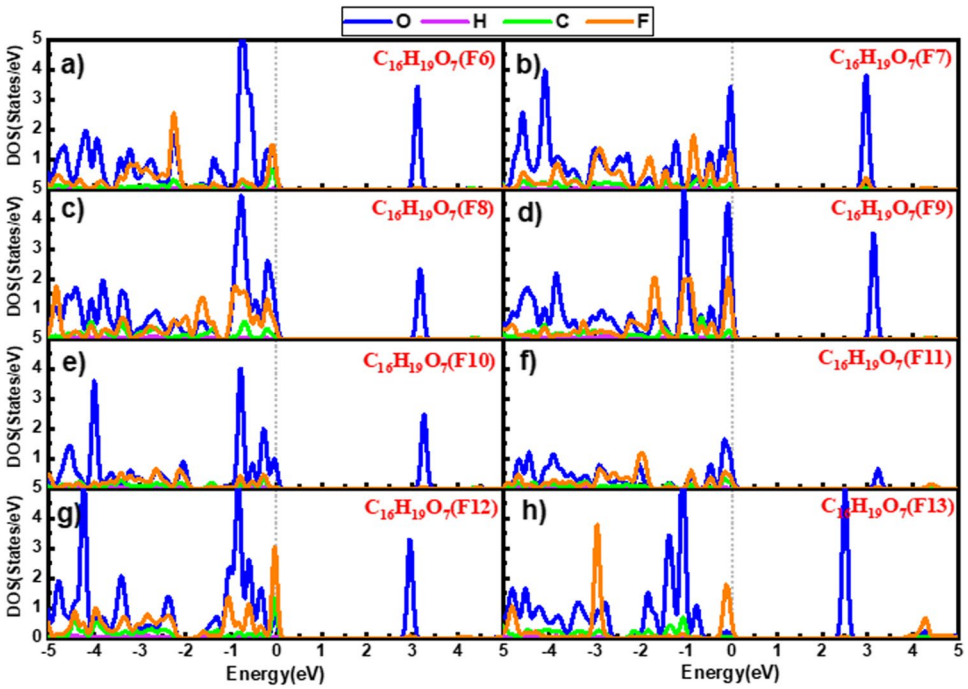

Fig. 5 The calculated total density of states for **a** $C_{16}H_{20}O_7$(F6), **b** $C_{16}H_{19}O_7$(F7), **c** $C_{16}H_{19}O_7$(F8), **d** $C_{16}H_{19}O_7$(F9), **e** $C_{16}H_{19}O_7$(F10), **f** $C_{16}H_{19}O_7$(F11), **g** $C_{16}H_{19}O_7$(F12), and **h** $C_{16}H_{19}O_7$(F13) at B category of compounds. In these compounds, the fluorine atom is bonded to carbon inside benzene ring. The gray dot line shows Fermi energy level

stability. The B category of fluorinated derivatives with the high HOMO–LUMO band gap are stable and hence are chemically harder than the A category of compounds. The B category compounds with the wide bandgap have semiconducting properties.

Calculated partial density of states of the pure and fluorinated-$C_{16}H_{20}O_7$ are shown in Fig. 4. It is seen that the contribution of s-states is small and p-states have a remarkable contribution. The s-states of hydrogen atom and p-states of carbon atom have a small contribution in the total density

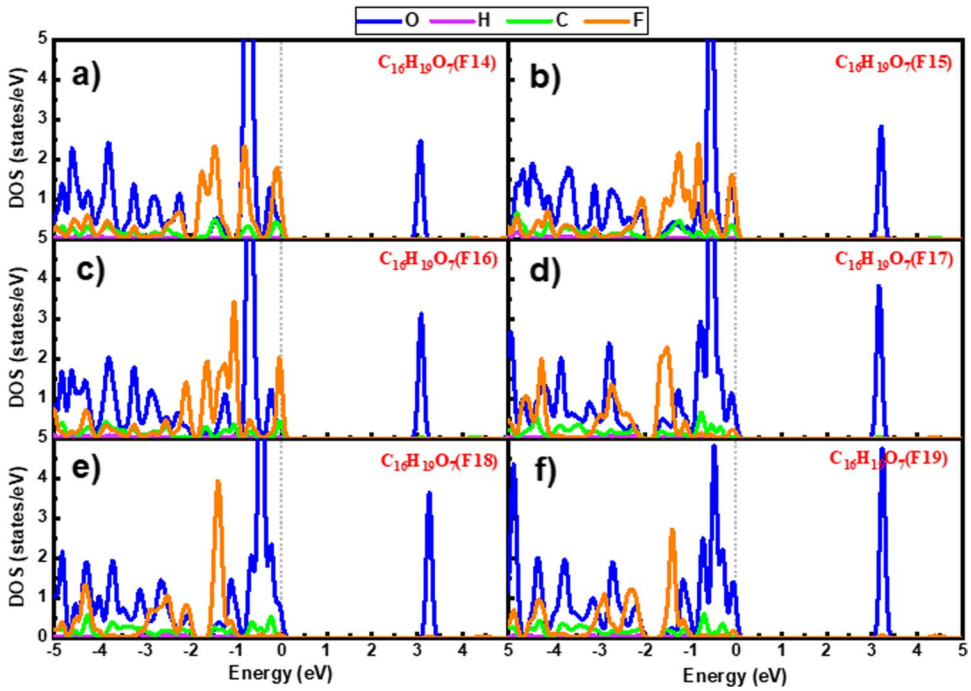

Fig. 6 The calculated total density of states for **a** $C_{16}H_{20}O_7$(F14), **b** $C_{16}H_{19}O_7$(F15), **c** $C_{16}H_{19}O_7$(F16), **d** $C_{16}H_{19}O_7$(F17), **e** $C_{16}H_{19}O_7$(F18), and **f** $C_{16}H_{19}O_7$(F19) at B category of compounds. In these compounds, the fluorine atom is bonded to carbon outside benzene ring with three hydrogen bonds. The gray dot line shows Fermi energy level

Table 1 Obtained electrical and optical data for $C_{16}H_{20}O_7$ and $C_{16}H_{19}O_7$(F1) to $C_{16}H_{19}O_7$(F19)

Compound	E_g	ε_0^x	ε_0^y	ε_0^z	n_0^x	n_0^y	n_0^z
$C_{16}H_{20}O_7$	3.108 Γ-Γ	2.55426	3.12551	2.76888	1.59821	1.76792	1.664
$C_{16}H_{19}O_7$(F1)	0.32 Γ-Γ	2.93709	5.56767	3.71813	1.7138	2.3599	1.92831
$C_{16}H_{19}O_7$(F2)	0.1 Γ-Γ	3.66598	4.81036	4.25495	1.91499	2.19384	2.06315
$C_{16}H_{19}O_7$(F3)	0.08 Γ-Γ	3.78383	4.31200	3.58904	1.94575	2.07696	1.89469
$C_{16}H_{19}O_7$(F4)	0.2 Γ-X	6.53135	4.09997	4.71215	2.55698	2.02496	2.17119
$C_{16}H_{19}O_7$(F5)	1.367 X-X	2.82586	3.63965	2.97702	1.70836	1.94341	1.75802
$C_{16}H_{19}O_7$(F6)	3.037 Γ-Γ	2.57459	3.03608	2.79362	1.60455	1.74244	1.67142
$C_{16}H_{19}O_7$(F7)	2.886 Γ-Γ	2.56741	3.08806	2.88624	1.60232	1.75729	1.6989
$C_{16}H_{19}O_7$(F8)	3.102 Γ-Γ	2.57263	3.08293	2.86346	1.60395	1.75583	1.69218
$C_{16}H_{19}O_7$(F9)	3.075 Γ-Γ	2.57579	3.00751	2.82737	1.60493	1.73422	1.68148
$C_{16}H_{19}O_7$(F10)	3.168 Γ-Γ	2.56482	3.04579	2.84275	1.60152	1.74525	1.68607
$C_{16}H_{19}O_7$(F11)	3.144 Γ-X	2.53929	3.08661	2.79723	1.59352	1.75688	1.6725
$C_{16}H_{19}O_7$(F12)	2.883 Γ-Γ	2.65766	3.18483	2.96684	1.63024	1.78462	1.72246
$C_{16}H_{19}O_7$(F13)	2.431 Γ-Γ	2.60264	3.27018	2.87743	1.61327	1.80837	1.6963
$C_{16}H_{19}O_7$(F14)	2.991 Γ-Γ	2.57207	3.03811	2.80066	1.60377	1.74302	1.67352
$C_{16}H_{19}O_7$(F15)	3.141 Γ-Γ	2.58228	3.03303	2.86712	1.60695	1.74156	1.69326
$C_{16}H_{19}O_7$(F16)	3.016 Γ-Γ	2.56321	3.06124	2.83321	1.60101	1.74964	1.68322
$C_{16}H_{19}O_7$(F17)	3.124 Γ-Γ	2.59432	3.06465	2.84178	1.61069	1.75062	1.68576
$C_{16}H_{19}O_7$(F18)	3.190 Γ-Γ	2.55828	3.04666	2.83598	1.59947	1.74547	1.68404
$C_{16}H_{19}O_7$(F19)	3.200 Γ-Γ	2.57751	3.00706	2.82623	1.60546	1.73409	1.68114

of states. In addition, it is shown for pure $C_{16}H_{20}O_7$, the p-states of the oxygen atom have a considerable contribution in HOMO and LUMO. In A category compounds by substitution of fluorine atom, the p-states of the fluorine atom have a remarkable contribution at the Fermi level and LUMO. On the other hand, in B category compounds such as $C_{16}H_{19}O_7$(F6 to F13), the p-states of the oxygen atom play a key role at LUMO (Fig. 5).

Figure 6 shows that the p-states of the fluorine atom have the high contribution near the Fermi level at HOMO and

Table 2 Calculated biological properties of compounds by Molsoft tool for the Lipinski's rule

	Molecular formula	Molecular weight	Number of hydrogen bond acceptor	Number of hydrogen bond donor	MolLogP	MolLogS	MolPSA(A^2)	MolVol(A^3)	pKa	BBB Score	Number of stereo centers	Drug- likeness model score
	$C_{16}H_{20}O_7$	324.12	7	5	−4.34	−0.71	99.17	314.30	< 0. /8.25	2.40	6	0.45
1	$C_{16}H_{19}O_7$(F1)	344.13	7	5	−0.19	−0.59	93.32	313.01	< 0. /10.03	2.40	6	0.04
2	$C_{16}H_{19}O_7$(F2)	344.13	7	5	−3.97	−0.51	93.55	313.68	< 0. /7.95	2.46	7	0.14
3	$C_{16}H_{19}O_7$(F3)	344.13	7	5	−3.95	−0.58	93.60	314.88	< 0. /7.95	2.46	7	0.25
4	$C_{16}H_{19}O_7$(F4)	344.13	7	5	−3.93	−0.58	93.60	313.67	< 0. /7.95	2.46	7	0.24
5	$C_{16}H_{19}O_7$(F5)	344.13	7	5	−0.62	−0.60	93.25	321.15	< 0. /13.35	2.46	7	−0.10
6	$C_{16}H_{19}O_7$(F6)	344.13	7	5	−4.10	−0.57	101.15	325.29	< 0.7.95	2.35	7	0.29
7	$C_{16}H_{19}O_7$(F7)	344.13	7	5	−4.13	−0.51	93.32	313.01	< 0. /7.95	2.46	7	0.04
8	$C_{16}H_{19}O_7$(F8)	344.13	7	6	−4.66	−0.47	101.43	311.75	< 0. /7.95	2.35	8	−0.24
9	$C_{16}H_{19}O_7$(F9)	344.13	7	6	−4.66	−0.47	101.43	311.75	< 0. /7.95	2.35	8	−0.24
10	$C_{16}H_{19}O_7$(F10)	344.13	7	6	−4.41	−0.59	101.15	325.91	< 0. /7.95	2.35	7	−0.12
11	$C_{16}H_{19}O_7$(F11)	344.13	7	6	−4.41	−0.53	101.15	331.45	< 0. /7.82	2.35	7	0.32
12	$C_{16}H_{19}O_7$(F12)	344.13	7	6	−4.50	−0.53	101.53	329.00	< 0. /7.95	2.35	7	−0.31
13	$C_{16}H_{19}O_7$(F13)	344.13	7	6	−3.88	−0.58	100.55	318.86	< 0. /6.75	2.36	7	0.24
14	$C_{16}H_{19}O_7$(F14)	344.13	7	6	−4.58	−0.50	101.68	317.31	< 0. /7.95	2.34	7	0.49
15	$C_{16}H_{19}O_7$(F15)	344.13	7	6	−4.58	−0.50	101.68	317.31	< 0. /7.95	2.34	7	0.49
16	$C_{16}H_{19}O_7$(F16)	344.13	7	6	−4.58	−0.50	101.68	317.31	< 0. /7.95	2.34	7	0.49
17	$C_{16}H_{19}O_7$(F17)	344.13	7	6	−4.42	−0.50	100.09	318.56	< 0. /7.95	2.37	7	0.35
18	$C_{16}H_{19}O_7$(F18)	344.13	7	5	−4.42	−0.50	100.09	318.56	< 0. /7.95	2.37	7	0.35
19	$C_{16}H_{19}O_7$(F19)	344.13	7	5	−4.42	−0.50	100.09	318.56	< 0. /7.95	2.37	7	0.35

Table 3 Chemical name, formula, and inhibition constant (Ki), binding energy, the number of hydrogen bond between ligand and HIV protease (HB), the names of residue that participated in hydrogen bonds

HIV protease							
Name	Formula	Ki (ℳM)	Binding energy (kcal/mol)	HB	Amino acids	Hydrophobic interactions	Amino Acids
1	$C_{16}H_{20}O_7$	27.32	− 6.23	10	GLN2B, THR4B (3), TRP6B, LEU10B (2), THR12B, GLU21B (2)	4	PRO1B, ILE3B, LEU10B, THR12B
4	$C_{16}H_{19}O_7$(F4)	26.9	− 6.23	10	GLN7B (2), THR4B (3), TRP6B, LEU10B (2), THR12B, GLU21B	4	PRO1B, ILE3B, LEU10B, THR12B
6	$C_{16}H_{19}O_7$(F6)	54.53	− 5.82	6	THR4A (2), GLN7A (2), ARG8A (2)	2	TRP6A, LEU10A
7	$C_{16}H_{19}O_7$(F7)	15	− 6.58	9	GLN2B (2), THR4B (2), TRP6B, LEU10B (2), THR12B, GLU21B	4	PRO1B, ILE3B, LEU10B, THR12B
8	$C_{16}H_{19}O_7$(F8)	51.88	− 5.85	10	TRP6B, GLN2B, THR4B (3), LEU10B (2), THR12B, GLU21B (2)	4	PRO1B, ILE3B, LEU10B, THR12B
9	$C_{16}H_{19}O_7$(F9)	112.95	− 5.38	10	THR12B, GLN2B, THR4B (3), TRP6B, LEU10B (2), GLU21B (2)	4	PRO1B, ILEU3B, LEU10B, THR12B
10	$C_{16}H_{19}O_7$(F10)	36.96	− 6.05	7	THR4A, TRP6A, GLN7A, ARG8A (3), LEU10A	4	TRP6A, GLN7A, ARG8A, LEU10A
11	$C_{16}H_{19}O_7$(F11)	37.32	− 6.04	11	GLN2B (2), THR4B (3), TRP6B, LEU10B (2), THR12B, GLU21B (2)	4	PRO1B, ILE3B, LEU10B, THR12B
12	$C_{16}H_{19}O_7$(F12)	31.72	− 6.14	10	GLN2B (2), THR4B (3), TRP6B, LEU10B (2), THR12B, GLU21B	4	PRO1B, ILE3B, LEU10B, THR12B
13	$C_{16}H_{19}O_7$(F13)	33.76	− 6.1	11	GLN2B (2), THR4B (3), TRP6B, LEU10B (2), THR12B, GLU21B (2)	3	ILE3B, LEU10B, THR12B
14	$C_{16}H_{19}O_7$ (F14)	122.34	− 5.34	9	GLN2B, THR4B (2), TRP6B, LEU10B, THR12B, GLU21B (3)	3	PRO1B, ILE3B, THR12B
15	$C_{16}H_{19}O_7$(F15)	83.39	− 5.56	11	GLN2B (2), THR4B (3), TRP6B, LEU10B (2), THR12B, GLU21B (2)	3	PRO1B, ILE3B, THR12B
16	$C_{16}H_{19}O_7$(F16)	684.75	− 4.32	8	ARG8A (4), GLN7A (3), LEU10A	1	LEU10A
17	$C_{16}H_{19}O_7$(F17)	63.5	− 5.73	10	GLN2B (2), THR4B (2), TRP6B, LEU10B (2), THR12B, GLU21B (2)	4	PRO1B, ILE3B, LEU10B, THR12B
18	$C_{16}H_{19}O_7$(F18)	33.94	− 6.1	11	GLN2B (2), THR4B (3), TRP6B, LEU10B (2), THR12B, GLU21B (2)	4	PRO1B, ILE3B, LEU10B, THR12B
19	$C_{16}H_{19}O_7$ (F19)	30.93	− 6.15	2	LEU10, THR12		

the contribution of the p-states of the oxygen atom is small. Nevertheless, the Tables 2 and 3 indicate that the p-states of the fluorine and oxygen atoms have the main role in the reactivity of $C_{16}H_{19}O_7$(F14) and $C_{16}H_{19}O_7$(F16) compounds with COVID-19 and HIV proteases, respectively.

For the other $C_{16}H_{19}O_7$(F17), $C_{16}H_{19}O_7$(F18), and $C_{16}H_{19}O_7$(F19) compounds, the p-states of the oxygen atom originates near the Fermi level and the contribution of the p-states of the fluorine is ignorable and it is observed in the semi-core region. Hence, in $C_{16}H_{19}O_7$(F17) compound, it is clear that the p-states of the oxygen are affected in reactivity with HTLV protease. In other words, in some of the B category compounds $C_{16}H_{19}O_7$(F14 to F19), the p-states of the oxygen and fluorine atoms play a key role in the pharmaceutical properties of compounds.

To analyze the optical properties of a material, we must first obtain its dielectric function, which consists

of both real and imaginary parts ($\varepsilon_1(\omega)$, $\varepsilon_2(\omega)$) [30]. It is the linear response of a system to electromagnetic waves. This function is calculated by considering the transition between the occupied and unoccupied states in the energy band or the transition between HOMO and LUMO states. The imaginary and real parts of the dielectric function are defined as [31]:

$$\varepsilon_1(\omega) = \delta_{ij} + \frac{2}{\pi} P \int_0^{\infty} \frac{\omega' \varepsilon_2(\omega')}{\omega'^2 - \omega^2} d\omega'. \quad (1)$$

$$\varepsilon_2(\omega) = \frac{4\pi e^2}{m^2\omega^2} \sum_{c,v} \int dk < c_k | P^a | v_k >< v_k | P^a | c_k > \delta(\varepsilon_{c_k} - \varepsilon_{v_k} - \omega) \quad (2)$$

where P represents the Cauchy integral part. $|c_k\rangle$ and $|v_k\rangle$ are the electronic states at the conduction band and the valence band, respectively. In Fig. 7a and b, the real and imaginary parts of the dielectric function for some selected compounds such as $C_{16}H_{20}O_7$, $C_{16}H_{19}O_7$(F4), $C_{16}H_{19}O_7$(F6), $C_{16}H_{19}O_7$(F13), $C_{16}H_{19}O_7$(F16), and $C_{16}H_{19}O_7$(F18) are plotted at x, y, and z directions. It is also seen from the $\varepsilon_1(\omega)$ spectra that obtained values are positive for all photon energy ranges (Fig. 7a). Hence, the dielectric nature of these compounds is identified. As seen, calculated $\varepsilon_{1x}(\omega)$ of all compounds have small variations between 2 and 4, except $C_{16}H_{19}O_7$(F4), while $\varepsilon_{1y}(\omega)$ and $\varepsilon_{1z}(\omega)$ vary between 0 and 6. So, there is an anisotropy between three directions of the $\varepsilon_1(\omega)$ spectrum corresponding to different polarizations. This polarization is due to the response of material to the electrical field of electromagnetic wave. The maximum value of polarized $\varepsilon_{1x}(\omega)$, $\varepsilon_{1y}(\omega)$, $\varepsilon_{1z}(\omega)$ for the $C_{16}H_{20}O_7$ and selected B category of compounds occurred at 3.1 eV. The first peaks are related to maximum values of variations $\varepsilon_2(\omega)$ spectrum (Fig. 7b). For $C_{16}H_{20}O_7$ and selected B category of compounds, the maximum value of $\varepsilon_2(\omega)$ is related to the transitions between the O-p state on HOMO and LUMO.

For $C_{16}H_{19}O_7$(F4) in A category of compounds, a maximum polarization of $\varepsilon_1(\omega)$ spectra is observed at the x-direction. The first peak of $\varepsilon_{2x}(\omega)$, $\varepsilon_{2y}(\omega)$, and $\varepsilon_{2z}(\omega)$ spectra is occurred around 1 eV due to the electronic transitions between the F-p state in HOMO and the O-p state in LUMO. For $C_{16}H_{19}O_7$(F13) compound, the first peaks of $\varepsilon_{1y}(\omega)$ and the $\varepsilon_{2y}(\omega)$ spectra at energy rang of 0–3 eV appears due to the small bandgap of compound and electronic transition from valence band to the unoccupied states in the conduction band. The calculated static dielectric constants at zero energy $\varepsilon_1(0)$ are presented in Table 1.

For B category of compounds, the maximum static dielectric constant $\varepsilon_1(0)$ is observed at the y-direction, while for $C_{16}H_{19}O_7$(F4) compound, the maximum value of the $\varepsilon_1(0)$ is obtained at the x-direction and it is about 6.35. It is found that the fluorination at B category of compounds does not effect on $\varepsilon_1(0)$ of compounds. As a result, a more detailed analysis showed that these

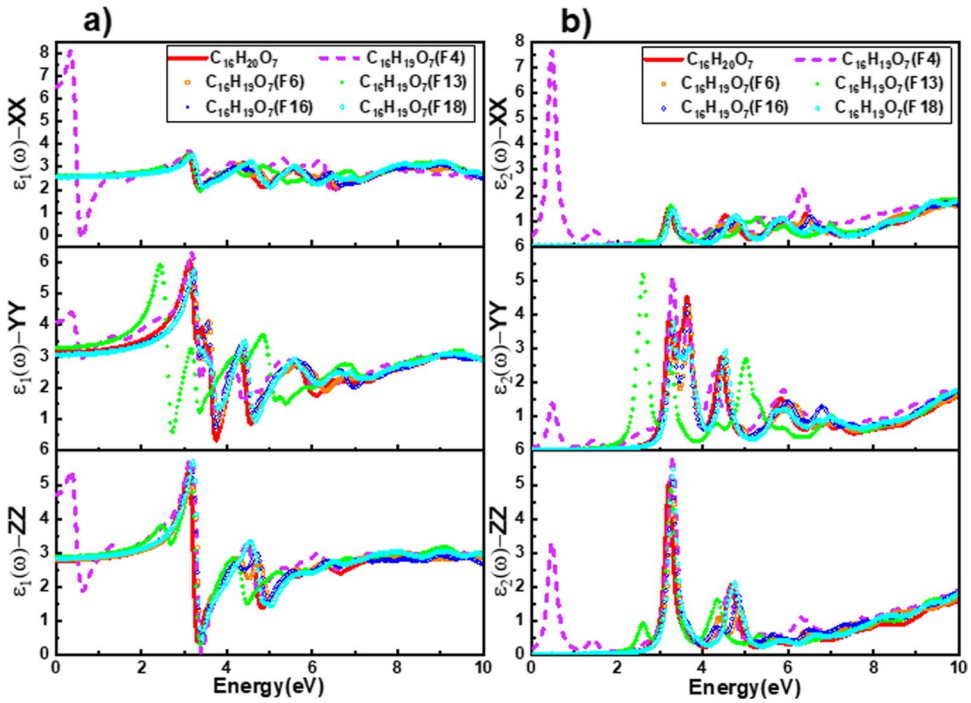

Fig. 7 **a** The real part of dielectric spectrum in terms of energy at three directions: x, y, z. **b** The imaginary parts of dielectric spectrum at x, y, and z directions

compounds exhibit the same mechanism of reaction between the drug molecules and the diseases proteases.

Electron energy-loss spectroscopy (EELS) is an analytical technique that evaluates the losing energy of an electronic beam during interaction with the substance. EELS corresponded to the HOMO and LUMO transitions of electrons and their elastic scattering [30]. It is an optical spectrum [32] used to get Plasmon energy. EELS is applied to identify [33] the different (organic) compounds used in biological systems or pharmaceuticals and to detect any impurities present [34] even during the synthesis or formulation process [35]. Also, EELS is used as a technique for quantification of penetration to cell walls [36]. EELS defined as:

$$ELOSS = -Im\left(\frac{1}{\varepsilon_1 + i\varepsilon_2}\right) = \frac{\varepsilon_2}{\varepsilon_1^2 + \varepsilon_2^2}. \quad (3)$$

Calculated EELS spectra for $C_{16}H_{20}O_7$, $C_{16}H_{19}O_7$(F4), $C_{16}H_{19}O_7$(F6), $C_{16}H_{19}O_7$(F13), $C_{16}H_{19}O_7$(F16), and $C_{16}H_{19}O_7$(F18) compounds are shown in Fig. 8 at the x-, y-, and the z-directions. According to this figure, the prominent peaks in ELOSS spectra are located at the energy of 24 eV for pure and fluorinated compounds of B category. The maximum values of peaks are about 2.3, 2.4, and 2.2 at the x-, y-, and the z-direction, respectively. The ELOSS spectrum for B category of compounds is the same as pure (Fig. 8b). The value of main peak of $C_{16}H_{19}O_7$(F4) is 1.9 at energy of 26 eV.

Absorption spectroscopy [37] is defined as the measurement of the interaction between electromagnetic radiation, and a substance is another important analytical chemistry technique to ensure the safety, quality, and efficacy of compounds used in the pharmaceutical industry. Especially, ultraviolet/visible (UV/Vis) light spectroscopy can be used in the manufacturing process to identify contaminants within a substance or measure the kinetics of a reaction. In this study, we have investigated the absorbance spectra of pure and fluorinated $C_{16}H_{20}O_7$ compounds, as in the following:

$$\alpha(\omega) = \frac{\sqrt{2\omega}}{c}\left\{\left[\varepsilon_1^2(\omega) + \varepsilon_2^2(\omega)\right]^{\frac{1}{2}} - \varepsilon_1(\omega)\right\}^{\frac{1}{2}}. \quad (4)$$

Figure 9 displays the absorbance spectra of the pure and selected fluorinated compounds at the x-, y-, and the z-directions. The maximum absorption is in the wavelength range of 50 to 125 nm that is at the extreme ultraviolet radiation (EUV) region. As shown, all compounds have the considerably same behavior in these regions. In ultraviolet (UV) regions in the wavelength range of 200 nm to 400 nm, the absorption coefficients are less than the EUV region (inset panels of Fig. 4). Also, there are some slight differences among the absorbance spectra of compounds. High similarity in absorbance spectrum of fluorinated compounds with pure is crucial from the viewpoints of pharmaceutical usages.

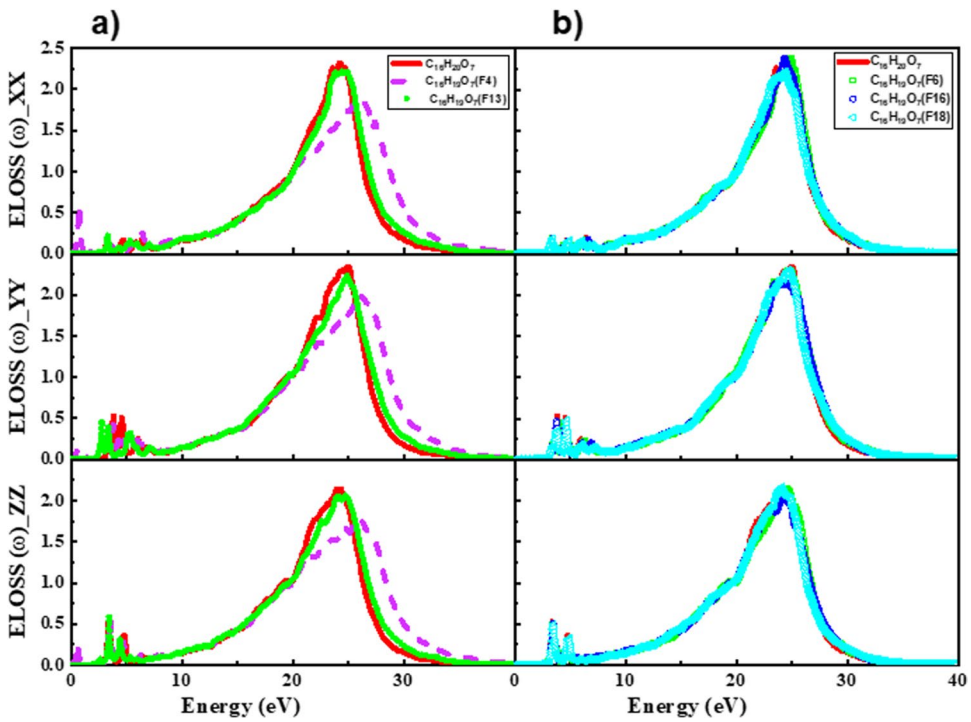

Fig. 8 The variation of energy electron loss spectrum L (ω) at three directions: x, y and z. **a** and **b** The comparison between $C_{16}H_{20}O_7$ and the other compounds with the small bandgap

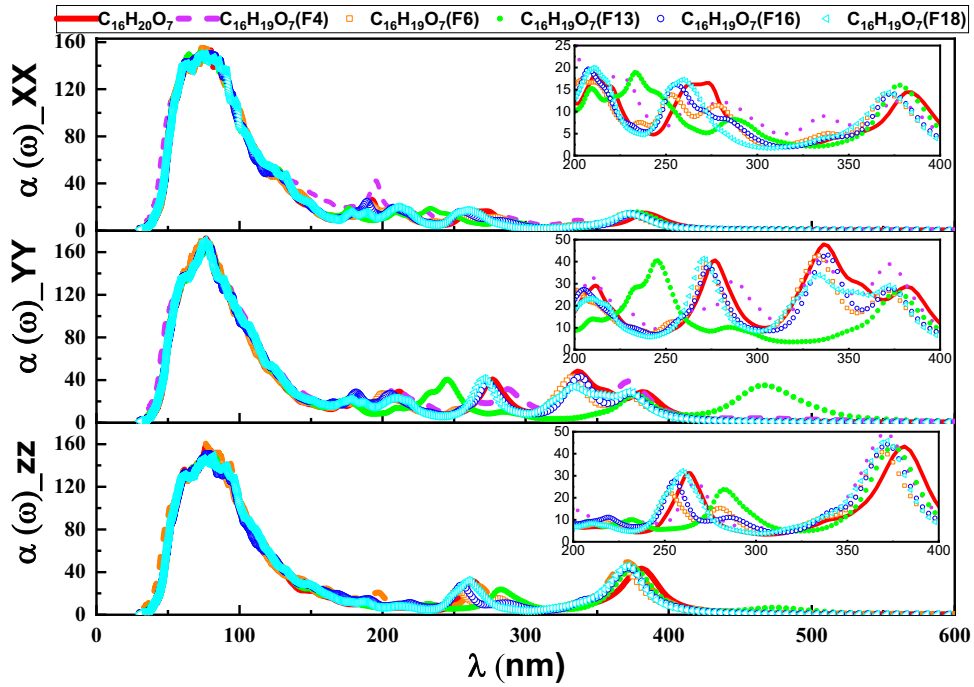

Fig. 9 Calculated absorption spectra at x, y, z directions. The inset plots are the absorption spectra in terms of wavelength at ultraviolet regions. The absorption spectra are in the unit of 10^4 cm.$^{-1}$

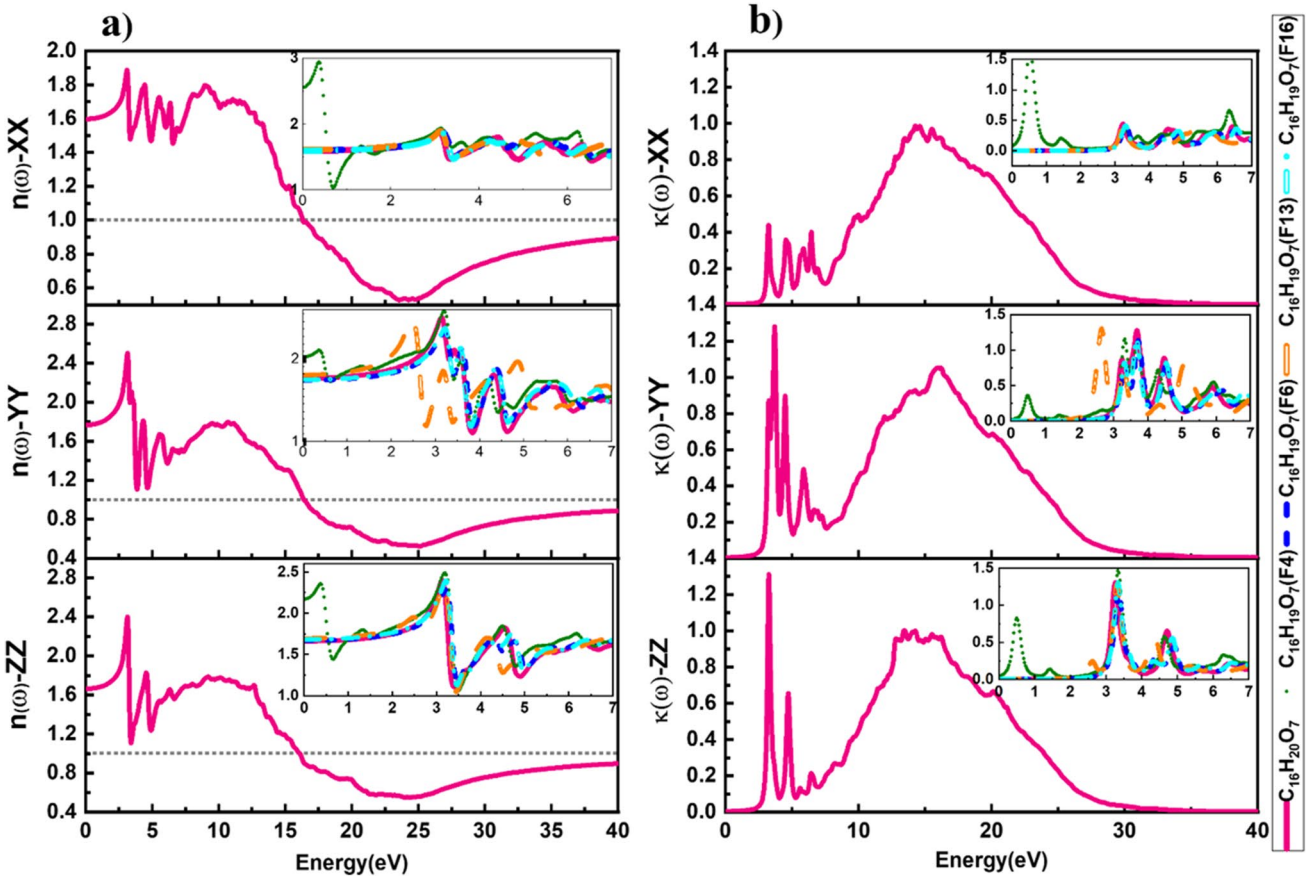

Fig. 10 **a** and **b** The calculated refraction index and extinction coefficient spectra for pure $C_{16}H_{20}O_7$ at x, y, z directions. The inset plots are related to the comparison of refraction index and extinction coefficient spectra among selected six compounds at the energy range 0 to 7 Ev

One of the most known physicochemical properties of substances is the refractive index [38]. The refractive index, that means the ratio of the speed of light passing in the vacuum to its passing through a substance, is complex quantity $\hat{n}(\omega) = n(\omega) + k(\omega)$, where $n(\omega)$ is the refractive index for the real part and $k(\omega)$ is the extinction coefficient for the imaginary part. We have calculated as follows:

$$n(\omega) = \frac{1}{\sqrt{2}}\sqrt{\sqrt{\varepsilon_1^2(\omega) + \varepsilon_2^2(\omega)} + \varepsilon_1(\omega)} \tag{5}$$

$$k(\omega) = \frac{1}{\sqrt{2}}\sqrt{\sqrt{\varepsilon_1^2(\omega) + \varepsilon_2^2(\omega)} - \varepsilon_1(\omega)} \tag{6}$$

Refractive index is mainly applied to identify a particular substance, confirm its purity, or measure its concentration. Generally, it is used to measure the concentration of a solute in an aqueous solution. So, it has the most applications to the determination of drug concentration in the pharmaceutical industry.

We have presented the refractive index and extinction coefficient spectra in terms of photon energy for $C_{16}H_{20}O_7$ compound, as shown in Fig. 10a and b. The calculated $n(\omega)$and$k(\omega)$ spectra exhibit some anisotropic behavior at the energy range 0 to 7.5 eV. For high-energy photons (higher than 16 eV), $n(\omega)$ is lesser than unity which means that the nature of the compound changes from linear to non-linear, and the superluminal effect occurs. The inset graphs are related to comparing of refraction index and extinction coefficient spectra among selected compounds at energy range 0 to 7 eV. Generally, it is shown that the refractive index and extinction coefficient spectra have the same behaviors for the pure and B category of compounds, except $C_{16}H_{19}O_7$(F13) compound at the y-direction. The calculated static refractive index at the zero energy is given in Table 1. According to Table 1, for the pure $C_{16}H_{20}O_7$ and B category of compounds, the maximum value of the static refractive index is about 1.7 at the y-direction. The highest value of static refractive index is related to the $C_{16}H_{19}O_7$(F4) compound at the x-direction.

Finally, we have calculated the reflectivity spectra R(ω) as follows:

$$R(\omega) = \frac{(n-1)^2 + k^2}{(n+1)^2 + k^2}, \tag{7}$$

where n and k are refractive index and extinction coefficient, respectively. As seen in Fig. 11, the main peak about 0.33 at 0.5 eV, is related to the $C_{16}H_{19}O_7$(F4) compound from category A of compounds at the x-direction. Other compounds have not considerable reflectivity at the x-direction. For the other compounds, the maximum value of reflectivity of 0.3 at 3.25 eV is observed at the z-direction. For pure $C_{16}H_{20}O_7$ compound, large optical reflectivity anisotropy is observed at 6.5 eV in the x-, y-, z-directions. On the other hand, the behaviors of reflectivity spectra for the pure and the other B category of compounds, except $C_{16}H_{19}O_7$(F13), are quite similar at the y-direction.

In summary, in comparison with the experimental results, there is not the available experimental

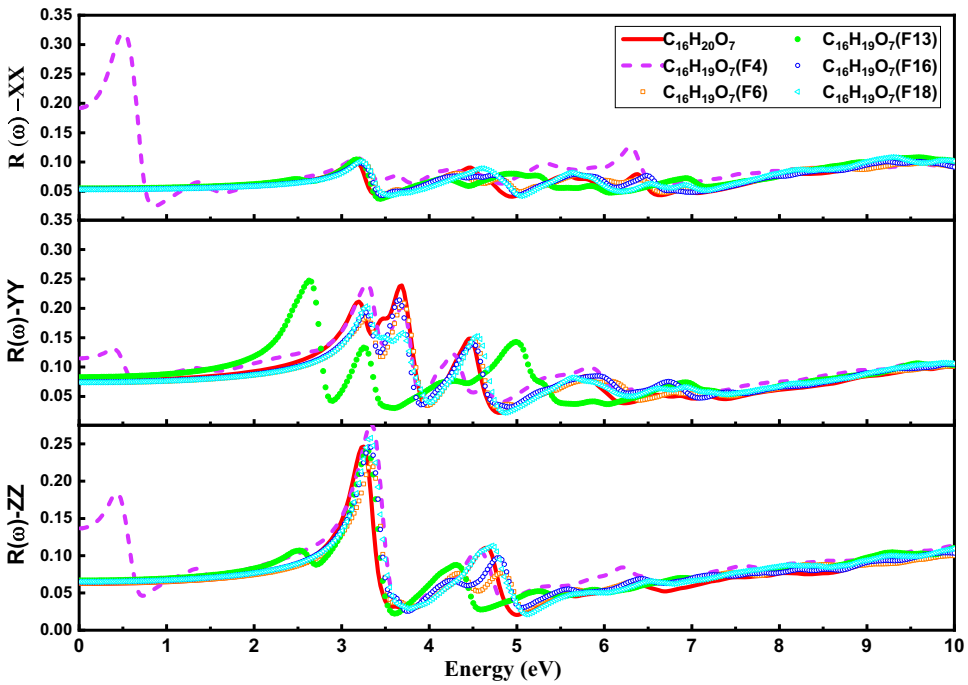

Fig. 11 The calculated reflectivity spectra at x, y, and z directions

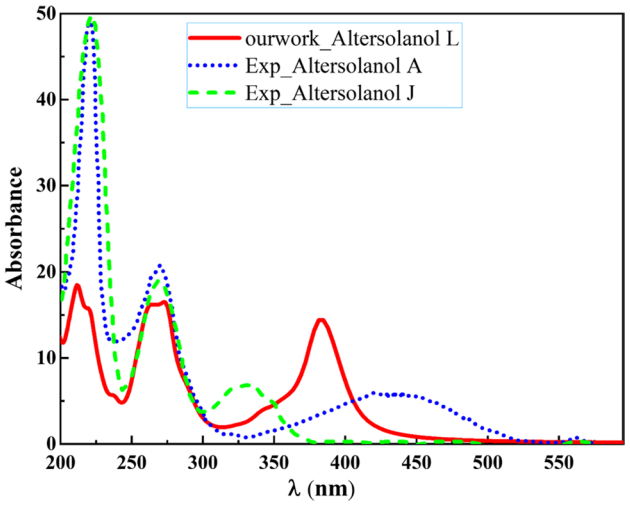

Fig. 12 Comparison of calculated absorption spectrum Altersolanol L at the x-direction (in unit of 10^{-4}/Cm) with the experimental absorption spectra of Altersolanol A and Altersolanol J (without unit) [39]

evidence regarding the obtained results. However, there are fewer studies for absorption spectra by Hassan like Altersolanol A and Altersolanol J compounds [39] providing stronger evidence that it makes these materials suitable for pharmaceutical applications. In Fig. 12, we have compared obtained absorption spectrum for Altersolanol L with the experimental results for the same family of compounds for Altersolanol A and Altersolanol J compounds [39]. There is a good agreement between obtained absorption spectra at 224 and 275 nm with the other experimental results.

B. Prediction of drug-likeness

To predict drug-likeness of designed compounds, Molsoft software was used to estimate the pharmacological similarity of fluorinated derivatives with the five famous Lipinski's rules [40]. Lipinski's rule of five: As per Lipinski's rule of five [40], an orally active drug has (i) not more than five hydrogen bond donors (OH and NH groups), (ii) not more than ten hydrogen bond acceptors, (iii) a molecular weight under 500, and (iv) a partition coefficient log P under 5.

The capacity of the hydrogen bond is an essential factor in the permeability of the drug to the body [41]. Drug permeation or absorption in the body decreases with increasing the number of hydrogen bond donors to more than five and the number of hydrogen bond acceptors to more than 10. The designed compounds have 7 H-bond donors and 5 to 7 H-bond acceptors (Table-2). Lipophilicity and partition coefficient (log P) values are an indicator of the tendency of the compound to dissolve in the membrane. The values describe the uptake, distribution, biotransformation, and excretion of organic chemicals in biological systems [42]. High log P values represent high solubility and good penetration of lipid membranes but low solubility in water and cytoplasm. Log P values for $C_{16}H_{20}O_7$ and its fluorinated derivatives are approximately – 4 (Table 2) confirming higher affinity to the aqueous phase.

Blood–brain barrier (BBB) is a physicochemical barrier that maintains homeostasis of the central nervous system (CNS) to regulate the content of the interstitial fluid in the brain [42]. The blood–brain barrier prediction score is greater than four indicates the chemical pass the BBB. This score was calculated in Molsoft server by Gupta

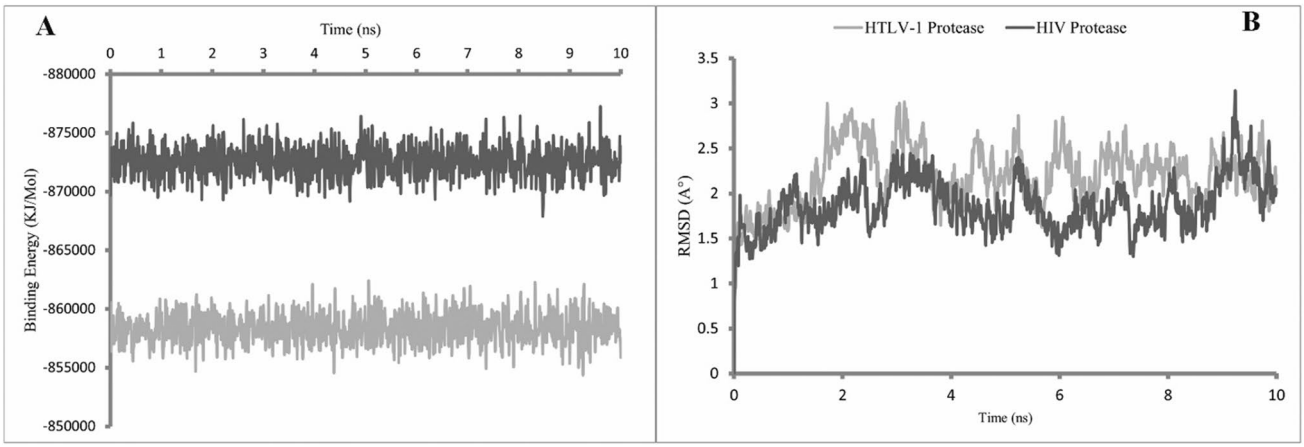

Fig. 13 RMSD and total energy for HTLV protease during 10 ns molecular dynamics simulations. **a** Root mean square deviations (RMSDs) of HTLV-1 protease (purple) and HIV protease (light purple). **b** Total energy of HTLV-1 protease (light purple) and HIV protease (purple). Both parameters confirm structural stability of protease during simulation

et al. [43]. In this study, the BBB scores for $C_{16}H_{20}O_7$ and its fluorinated derivatives are 2.3 to 2.4 (Table 2) confirming them as unpenetrated chemicals to the brain.

The aqueous solubility of compounds is significant to oral absorption and distribution characteristics. All designed compounds have aqueous solubilities of Mol-LogS from − 0.7 to − 0.5. Based on ADME prediction, the $C_{16}H_{19}O_7$(F14), $C_{16}H_{19}O_7$(F15), and $C_{16}H_{19}O_7$(F16) have the best favorite drug-likeness rather than $C_{16}H_{20}O_7$. In them, fluorine is substituted with hydrogen in the methyl group (C16). To get more optimal and reliable results from the docking process, first, the 3D structure of HTLV-1 protease (PDB codes: 3LIN) and HIV protease (PDB code: 1YT9) were optimized by Gromacs in 10 ns. The RMSD results (root mean square deviation) and total energy showed that the protein structure during running time is stable (Fig. 13). HTLV-1 and HIV have a homodimer activity structurally, and each monomer of these enzymes has 125 and 99 amino acids, respectively. The active site of these proteases is located in a valley between two monomers ASP25, and ASP32 in HIV and HTLV-1 protease, respectively, in each monomer act as the main catalytic amino acids [44]. The sequence alignment evidence shows 28% identity between the primary sequence of HTLV-1 protease and HIV protease enzyme, while they are more conserved in the active site (45%) [45].

Ruker et al. [46] investigated similarities and differences in HTLV and HIV protease-ligand interactions on the molecular level. They showed the main protease-ligand interactions happen in similar regions of proteases: the N-terminal (HIV-PR R8; HTLV-PR R10), the active site (HIV-PR L23-D30; HTLV-PR L30-D36), the flap region (HIV-PR M46-I50; HTLV-PR S55-A59), and the lateral loops of the ligand-binding site (HIV-PR P81-I84; HTLV-PR N97-I100). Also, they showed, based on interaction energy profile, however, some well-conserved residues have strong interactions with the substrate in two proteases (R8/R10, A28/A35, D29/D36, G49, G58, and I84/I100), there are many not-conserved residues have strong interactions in the similar ligand binding pockets in HIV and HTLV protease (D30/M37, I47/V56, G48/L57, I50/A59, and V82/W98).

The $C_{16}H_{20}O_7$ and fluorinated compounds interacted with residues in the N-terminal of HIV and HTLV protease that are listed in Tables 3 and 4 in detail. In designed fluorinated compounds, $C_{16}H_{19}O_7$(F14), $C_{16}H_{19}O_7$(F15), and $C_{16}H_{19}O_7$(F16) have the best drug-likeness other than else. The $C_{16}H_{19}O_7$(F16) interacted with ARG8A (4), GLN7A (3), and LEU10A via hydrogen bonds in complex with HIV protease and $C_{16}H_{19}O_7$ (F14) connected to ARG10B, GLU28B (2), and ALA99B in HTLV protease. Nowadays, the world has been confronted the novel coronavirus (2019-nCOV)–dependent acute respiratory illness. It seems that imbalanced immune cells' activity can lead to the self-destruction of the host, resulting in high severity of COVID-19. Lim et al. [47] showed that a load of the virus decreases significantly in Korean patients diagnosed with COVID-19 when treated with lopinavir and ritonavir as two potent inhibitors of HIV protease. Dipti Mothay et al. [48] analyzed some potent inhibitors of COVID-19 protease that they are able to treat the HIV patients. Based on in silico study, Arg141of protease model of COVID-19 had molecular interactions with lopinavir, ketoamide, and ritonavir; Gln175 showed molecular interactions with nelfinavir and Thr75 and His176 from COVID-19 protease bind to remdesivir via noncovalent interactions. Günther et al. [49] introduced an active site and two allosteric binding sites for potent inhibitors of COVID-19 protease: The hydrophobic pocket is created by Ile213, Leu253, Gln256, Val297, and Cys300 within the C-terminal dimerization domain; the deep groove is located between the catalytic domains (Cys145 and His 141) and the dimerization domain (Arg4, Ser10, Gly11, Glu14, Asn28, Ser139, Phe140, Ser147, Glu290, Arg298) [50]. Günther, et al. showed that AT7519 is the only compound in their screen that bound to loop 107–110, Val202, and Thr292, and the chlorinated benzene ring in AT7519known as a potent compound for human cancer treatment [49] is engaged in various van der Waals interactions to loop 107–110, Val202, and Thr292 [50, 51]. Muramatsu et al. introduced substrate-binding residues His41, Met49, Gly143, Ser144, His163, His164, Met165, Glu166, Leu167, Asp187, Arg188, Gln189, Thr190, Ala191, and Gln192 for COVID-19 protease [52]. Based on our results, $C_{16}H_{19}O_7$ (F14) has a similar binding mode to COVID-19 protease with the AT7519 compound mentioned above. Two compounds are halogenated, which may effect on binding affinity to the special site of the COVID-19 protease. This site may be a target to inhibitor binding for COVID-19 protease. The $C_{16}H_{19}O_6$ (F14) has hydrogen bonds and hydrophobic interactions with Arg298 in dimerization domain and loop 107–110. It may introduce as potent inhibitor to COVID-19 protease (Table 5).

Conclusions

First principle calculations based on DFT and molecular docking approaches are used for investigation of optoelectronic and biological properties of pure $C_{16}H_{20}O_7$ and substituted by fluorine. All compounds are classified into one of two groups: category A or category B. The electronic results show that the

Table 4 Chemical name, formula, and inhibition constant (Ki), binding energy, the number of hydrogen bond between ligand and HTLV- (HB), the names of residue that participated in hydrogen bonds

HTLV protease									
Name	Formula	Ki (𝑀M)	Binding energy (kcal/mol)	HB	Amino acids	Hydrophobic interactions	Amino Acids	Π-Stacking	Halogen bond
1	$C_{16}H_{20}O_7$	144.77	−5.24	6	ARG10A, VAL2A, PRO4A (2), VAL12A (2)	4	ILE3A (2), PRO4A, ASP6A		
4	$C_{16}H_{19}O_7$(F4)	32.53	−6.12	4	PRO4A (2), VAL12A (2)	3	ILE3A, PRO4A, ASP6A		
6	$C_{16}H_{19}O_7$(F6)	77.1	−5.61	4	ARG10B, GLU28B, ASN97B, ALA99B	3	LEU30B, ALA43B, ALA99B	TRP98B	
7	$C_{16}H_{19}O_7$(F7)	68.2	−5.68	4	ARG10B, GLU28B (2), ALA99B	2	GLU28B, ALA99B		
8	$C_{16}H_{19}O_7$(F8)	134.12	−5.28						
9	$C_{16}H_{19}O_7$(F9)	99.6	−5.46	6	ARG10B, GLU28B (3), ASN97B, ALA99B	2	LEU30B, TRP98B		
10	$C_{16}H_{19}O_7$(F10)	154.6	−5.2	8	ARG10B, GLU28B (3), LEU30B, ASN97B, ALA99B(2)	3	ILE27B, LEU30B, TRP98B		
11	$C_{16}H_{19}O_7$(F11)	122.06	−5.34	8	ARG10B, GLU28B (3), LEU30B, ASN97B, ALA99B	1	LEU30B		
12	$C_{16}H_{19}O_7$(F12)	84.55	−5.56	6	ARG10A, LEU57B, GLY58B, GLY60A, GLY60B, HIS66A	1	PHE67A	TRP98A	LEU57B
13	$C_{16}H_{19}O_7$(F13)	83.43	−5.56	6	GLU28B (3), ASN97B (2), ALA99B	4	27B (2), TRP98B, ALA99B		
14	$C_{16}H_{19}O_7$ (F14)	278.27	−4.85	4	ARG10B, GLU28B (2), ALA99B	2	GLU28B, ALA99B		
15	$C_{16}H_{19}O_7$(F15)	242.2	−4.93	3	ILE3A (2), PRO4A	6	VAL2A, PRO4A (2), ARG10A, VAL12A (2)		
16	$C_{16}H_{19}O_7$(F16)	79.08	−5.6	3	ILE3A (2), PRO4A	6	VAL2A, PRO4A (2), ARG10A, VAL12A (2)		
117	$C_{16}H_{19}O_7$(F17)	392.96	−4.65	5	ARG10A VAL2A, PRO4A (2), ARG10A, VAL12A	**2**	ILE3A, ASP6A		
118	$C_{16}H_{19}O_7$(F18)	306.29	−4.79	6	VAL2A, PRO4A (2), ARG10A, VAL12A (2)	2	ILE3A, ASP6A		
119	$C_{16}H_{19}O_7$(F19)	102.08	−5.44	5	PRO4A (2), ARG10A, VAL12A (2)	2	ILE3A, ASP6A		

pure $C_{16}H_{20}O_7$ compound has a direct bandgap about 3.1 eV. The substitution of F atom at A category decreases the electronic bandgap, while the bandgap value of B category compounds is very similar to the pure one except for $C_{16}H_{19}O_7$(F13) compound. The wide bandgap nature of pure $C_{16}H_{20}O_7$ and fluorinated derivatives show that these materials can be used in biosensors applications. F-O bonds in the B category of compounds such as $C_{16}H_{19}O_7$(F14), $C_{16}H_{19}O_7$(F16), and $C_{16}H_{19}O_7$(F17), have a considerable contribution on their stability and chemical reactivity. It is shown that fluorination at B category of compounds does not effect on intrinsic static dielectric constant of compounds. So, these compounds have the same nature in the reaction between the drug molecules and the diseases proteases. The anisotropic optical properties

Table 5 Chemical name, formula, and inhibition constant (Ki), binding energy, the number of hydrogen bond between ligand and COVID-19 (HB), the names of residue that participated in hydrogen bonds

COVID-19 protease									
Name	Formula	Ki (μM)	Binding Energy (kcal/ol)	HB	Amino Acids	Hydrophobic interactions	Amino acids	Π-Stacking	Alogen bod
1	$C_{16}H_{20}O_7$	144.77	−5.24	9	ARG4A (2), VAL2A, LYS5A(3), TRP207, LEU282, GLU288 (2)	3	TRP207, LEU282, GLU288		
4	$C_{16}H_{19}O_6$ (F4)	32.53	−6.12	5	GLN110, THR111, ILE152, ASP153, ASP295	3	ASN151, VAL-303ASP153		
6	$C_{16}H_{19}O_6$ (F6)	77.1	−5.61	6	PHE66(2), VAL68, ARG76, VAL77(2)	2	HIS64, ARG76		
7	$C_{16}H_{19}O_6$ (F7)	68.2	−5.68	6	PHE66, VAL68, LEU75, ARG76, VAL77(2)	2	HIS64, GLN74		
8	$C_{16}H_{19}O_6$(F8)	134.12	−5.28	9	TYR37(2), GLN83, LYS88, TYR101(3), PHE103(2)	4	TYR101, LYS102, PHE103(2)		
9	$C_{16}H_{19}O_6$ (F9)	99.6	−5.46	6	PHE66(2), VAL68, ARG76, VAL77(2)	2	HIS64, GLN74		
10	$C_{16}H_{19}O_6$ (F10)	154.6	−5.2	5	AASN63, PHE66(2), ARG76, VAL77	3	IASN63, HIS64, GLN74		
11	$C_{16}H_{19}O_6$ (F11)	122.06	−5.34	6	PHE66(3), VAL68, ARG76, VAL77	2	HIS64, LEU67		
12	$C_{16}H_{19}O_6$ (F12)	84.55	−5.56	5	VAL68, LEU75, ARG76, VAL77(2)	3	HIS64, LEU67, ARG76		
13	$C_{16}H_{19}O_6$ (F13)	83.43	−5.56	7	PHE66(2), VAL68, ARG76, VAL77(3)	1	HIS64		VAL68
14	C16H19O6 (F14)	278.27	−4.85	**7**	GLN110, THR111, ASN151, ASP153(2), THR292, ARG298	2	PHE294(2)		
15	$C_{16}H_{19}O_6$ (F15)	242.2	−4.93	5	VAL68, LEU75(2), ARG76, VAL77	2	GLN74, ARG76		
16	$C_{16}H_{19}O_6$ (F16)	79.08	−5.6	5	VAL68, LEU75, ARG76, VAL77 3A (2), PRO4A	2	GLN74, ARG76		
17	$C_{16}H_{19}O_6$ (F17)	392.96	−4.65	4	PHE305(3), ARG298	4	PRO9, ARG298, THR304, PHE305		
18	$C_{16}H_{19}O_6$ (F18)	306.29	−4.79	6	PHE66(2), VAL68, ARG76, VAL77(2)	2	HIS64, LEU67		
19	$C_{16}H_{19}O_6$ (F19)	102.08	−5.44	6	PHE66(2), VAL68, ARG76, VAL77(2)	2	HIS64, LEU67		

are observed in three directions. Because of the obtained optical spectra, the B category of compounds are predicted to be useful for different pharmaceutical usages in Covid-19, HIV, and HTLV proteases. Based on ADME prediction, the $C_{16}H_{19}O_7$(F14), $C_{16}H_{19}O_7$(F15), and $C_{16}H_{19}O_7$(F16) have the best favorite drug-likeness rather than pure $C_{16}H_{20}O_7$. The $C_{16}H_{19}O_6$ (F14) has hydrogen bonds and hydrophobic interactions with Arg298 in dimerization domain and loop 107–110. It may introduce as potent inhibitor to COVID-19 protease.

Acknowledgements The authors are grateful to P. Blaha (Vienna University of Technology, Austria) for his technical assistance using the WIEN2k code.

Author contribution All authors contributed to the study conception and design. M. Kheirabadi and H. A. Rahnamaye Aliabad designed

research and performed the simulations. Maliheh Azadparvar and M. Kheirabadi analyzed the data. The first draft of the manuscript was written by Maliheh Azadparvar and M. Kheirabadi and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Data availability Data will be available on reasonable request.

Code availability Not applicable.

Declarations

Conflict of interest The authors declare no competing interests.

References

1. Dadachova E, Bryan RA, Huang X, Moadel T, Schweitzer AD, Aisen P, Nosanchuk JD, Casadevall A (2007) Ionizing radiation changes the electronic properties of melanin and enhances the growth of melanized fungi. PLoS ONE 2(5):457. https://doi.org/10.1371/journal.pone.0000457
2. Raghukumar C, Raghukumar S (1998) Barotolerance of fungi isolated from deep-sea sediments of the Indian Ocean. Aquat Microb Ecol 15(2):153–163. https://doi.org/10.3354/ame015153
3. Sancho LG, de la Torre R, Horneck G, Ascaso C, de Los RA, Pintado A, Wierzchos J, Schuster M (2007) Lichens survive in space: results from the 2005 LICHENS experiment. Astrobiology 7(3):443–445. https://doi.org/10.1089/ast.2006.0046.PMID17630840.S2CID4121180
4. Jakubczyk D, Dussart F (2020) Selected Fungal Natural Products with Antimicrobial Properties. Molecules 25:911. https://doi.org/10.3390/molecules25040911
5. Linnakoski R, Reshamwala D, Veteli P, Cortina-Escribano M, Vanhanen H, Marjomäki V (2018) Antiviral Agents from Fungi: Diversity, Mechanisms and Potential Applications. Front Microbiol 9:2325. https://doi.org/10.3389/fmicb.2018.02325
6. Evidente A, Kornienko A, Cimmino A, Andolfi A, Lefranc F, Mathieu V, Kiss R (2014) Fungal metabolites with anticancer activity. Nat Prod Rep 31(5):617–627. https://doi.org/10.1039/c3np70078j
7. Adeleke BS, Babalola OO (2021) Pharmacological potential of fungal endophytes associated with medicinal plants: A review. Journal of Fungi 7(2):147. https://doi.org/10.3390/jof7020147
8. Dalinova AA, Salimova DR, Berestetskiy AO (2020) Fungi of the genera Alternaria as producers of biological active compounds and mycoherbicides. Appl Biochem Microbiol 56:256–272. https://doi.org/10.1134/S0003683820030023
9. Suemitsu R, Horiuchi K, Kubota M, Okamatsu T (1990) Production of alterporriols, altersolanols and macrosporin by Alternaria porri and A. Solani Phytochemistry 29(5):1509–1511. https://doi.org/10.1016/0031-9422(90)80110-3
10. Becker AM, Rickards RW, Schmalzl KJ, Yick HC (1978) Metabolites of Dactylaria lutea the structures of dactylariol and the antiprotozoal antibiotic dactylarin. J Antibiot 31(4):324–329. https://doi.org/10.7164/antibiotics.31.324
11. Evidente A, Rodeva R, Andolfi A, Stoyanova Z, Perrone C, Motta A (2011) Phytotoxic polyketides produced by Phomopsis foeniculi, a strain isolated from diseased Bulgarian fennel. Eur J Plant Pathol 130(2):173–182. https://doi.org/10.1007/s10658-011-9743-0
12. Assante G, Nasini G (1987) Identity of the phytotoxin stemphylin from Stemphylium botryosum with altersolanol A. Phytochemistry 26(3):703–705. https://doi.org/10.1016/S0031-9422(00)84768-8
13. Haraguchi H, Abo T, Fukuda A, Okamura N, Yagi A (1996) Mode of phytotoxic action of altersolanols. Phytochemistry 43(5):989–992. https://doi.org/10.1016/S0031-9422(96)00381-0
14. Zhang L, Lin D, Sun X, Curth U, Drosten C, Sauerherin L, Becke S, Rox K, Hilgenfel R (2020) Crystal structure of SARS-CoV-2 main protease provides a basis for design of improved α-ketoamide inhibitors. Science 368(6489):409–412. https://doi.org/10.1126/science.abb3405
15. Kuhnert M, Steuber H, Diederic WE (2014) Structural basis for HTLV-1 protease inhibition by the HIV-1 protease inhibitor indinavir. J Med Chem 57(14):6266–6272. https://doi.org/10.1021/jm500402c
16. Jbara AS, Othaman Z, Aliabad HA, Saeed MA (2017) Electronic and Optical Properties of γ-and-Alumina by First Principle Calculations. Adv Sci Eng Med 9(4):287–293. https://doi.org/10.1166/asem.2017.2007
17. Blaha P, Schwarz K, Madsen GKH, Kvasnicka D, Luitz J (2001) "wien2k". An augmented plane wave+ local orbitals program for calculating crystal properties. https://www.researchgate.net/publication/237132866
18. Pauling L (1960) The nature of the chemical bond. Cornell University Press, Ithaca. https://doi.org/10.1021/ja01355a027
19. O'Hagan D (2008) Understanding organofluorine chemistry. An introduction to the C–F bond. Chem Soc Rev 37(2):308–319. https://doi.org/10.1039/B711844A
20. Deng H, Ma L, Bandaranayaka N, Qin Z, Mann G, Kyeremeh K, Yi Yu, Shepherd T, Naismith JH, O'Hagan D (2014) Identification of fluorinases from Streptomyces sp MA37, Norcardia brasiliensis, and Actinoplanes sp N902–109 by genome mining. ChemBioChem 15(3):364–368. https://doi.org/10.1002/cbic.201300732
21. Murphy CD, (2003) New frontiers in biological halogenation. J Appl Microbiol 94:539–548. http://www1.udel.edu/chem/polenova/VHPO/Biological_Halogenation_New_Frontiers_JAppMicro2003.pdf
22. Johns K, Stead G (2000) Fluoroproducts—the extremophiles. J Fluorine Chem 104(1):5–18. https://doi.org/10.1016/S0022-1139(00)00251-7
23. Zheng CJ, Shao CL, Guo ZY, Chen JF, Deng DS, Yang KL, Chen YY, Fu XM, She ZG, Lin YC, Wang CY (2012) Bioactive hydroanthraquinones and anthraquinone dimers from a soft coral-derived Alternaria sp. fungus. J Nat Prod 75(2):189–97. https://doi.org/10.1021/np200766d
24. Hohenberg P, Kohn WJPR (1964) Density functional theory (DFT)." Phys. Rev 136 (1964): B864. https://th.fhi-berlin.mpg.de/meetings/DFT-workshop-2018/uploads/Meeting/ 08_Suhuai_Wei_LAPW.pdf
25. Froimowitz M (1993) HyperChem: a software package for computational chemistry and molecular modeling. Biotechniques 14(6):1010–1013. http://www.hypercubeusa.com/
26. WeiXing SHY, YouJin SHI (2004) Creating carbon nanotube and researching into its mechanics properties by HyperChem. Comput Appl Chem. https://doi.org/10.16866/j.com.app.chem2004.03.038
27. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ (2005) GROMACS: fast, flexible, and free. J Comput Chem 26(16):1701–1718. https://doi.org/10.1002/jcc.20291
28. Atilgan E, Hu J (2011) Improving protein docking using sustainable genetic algorithms. Int J Compu Inoform Sys Ind Manag App 3:248–255. https://cse.sc.edu/~jianjunh/paper/autodockx.pdf
29. DeLano WL (2002) The PyMOL user's manual. DeLano Scientific, San Carlos, CA, p 452. https://pymol.sourceforge.net/newman/userman.pdf
30. Bashi M, Aliabad HR, Mowlavi AA, Ahmad I (2017) 125Te NMR shielding and optoelectronic spectra in XTe3O8 (X= Ti, Zr, Sn and Hf) compounds: Ab initio calculations. J Mol Struct 1148:223–230. https://doi.org/10.1016/j.molstruc.2017.07.041

31. Aliabad HR, Rabbanifar S, Khalid M (2019) Structural, optoelectronic and thermoelectric properties of FeSb2 under pressure: bulk and monolayer. Physica B 570:100–109. https://doi.org/10.1016/j.physb.2019.06.001
32. Egerton RF (2008) Electron energy-loss spectroscopy in the TEM. Rep Prog Phys 72(1):016502. https://doi.org/10.1088/0034-4885/72/1/016502
33. Coxon PR, Chao Y, Horrocks BR, Gass M, Bangert U, Šiller L (2008) Electron energy loss spectroscopy on alkylated silicon nanocrystals. J Appl Phys 104(8):084318. https://doi.org/10.1063/1.3000566
34. Costa JL, Joy DC, Maher DM, Kirk KL, Hui SW (1978) Fluorinated molecule as a tracer: difluoroserotonin in human platelets mapped by electron energy-loss spectroscopy. Science 200(4341):537–539. https://doi.org/10.1126/science.644312
35. Das PP, Guzzinati G, Coll C, Gomez Perez A, Nicolopoulos S, Estrade S et al (2020) Reliable Characterization of Organic & Pharmaceutical Compounds with High Resolution Monochromated EEL Spectroscopy. Polymers 12(7):1434. https://doi.org/10.3390/polym12071434
36. Rapp AO, Bestgen H, Adam W, Peek RD (1999) Electron energy loss spectroscopy (EELS) for quantification of cell-wall penetration of a melamine resin. https://doi.org/10.1515/HF.1999.018
37. Pharmaceutical Chemistry: Drug Analysis Spectroscopic Methods (2020) Yerevan State Medical University. https://www.ysmubooks.am/uploads/handout_master_part_1_.pdf
38. Singh S (2002) Refractive index measurement and its applications. Physica Scripta 65(2):167. https://doi.org/10.1238/Physica.Regular.065a00167. & Mohan S, Kato E, Drennen III, JK, Anderson CA (2019) Refractive index measurement of pharmaceutical solids: a review of measurement methods and pharmaceutical applications. J Pharm Sci 108(11), 3478–3495. https://doi.org/10.1016/j.xphs.2019.06.029
39. Hassan AE (2007) Novel natural products from endophytic fungi of Egyptian medicinal plants: chemical and biological characterization (Doctoral dissertation, Düsseldorf, Univ., Diss. https://d-nb.info/98528207X/34
40. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ (1997) Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. Adv Drug Deliv Rev 23(1–3):3–25. https://doi.org/10.1016/S0169-409X(96)00423-1
41. Refsgaard HH, Jensen BF, Brockhoff PB, Padkjær SB, Guldbrandt M, Christensen MS (2005) In silico prediction of membrane permeability from calculated molecular parameters. J Med Chem 48(3):805–811. https://doi.org/10.1021/jm049661n
42. Zlolovic BV (2008) The blood-brain barrier in health and chronic neurodegenerative disorders. Neuron 57(2):178–201. https://doi.org/10.1016/j.neuron.2008.01.003
43. Gupta M, Lee HJ, Barden CJ, Weaver DF (2019) The blood–brain barrier (BBB) score. J Med Chem 62(21):9824–9836. https://doi.org/10.1021/acs.jmedchem.9b01220
44. Li C, Li X, Lu W (2010) Total chemical synthesis of human T-cell leukemia virus type 1 protease via native chemical ligation. Pept Sci 94:487–494. https://doi.org/10.1002/bip.21375
45. Li M, Laco GS, Jaskolski M, Rozycki J, Alexandratos J, Wlodawer A, Gustchina A (2005) Crystal structure of human T cell leukemia virus protease, a novel target for anticancer drug design. Proc Natl Acad Sci 102(51):18332–18337. https://doi.org/10.1073/pnas.0509335102
46. Rücker P, Horn AHC, Meiselbach H, Sticht H (2011) A comparative study of HIV-1 and HTLV-I protease structure and dynamics reveals a conserved residue interaction network. J Mol Model 17(10):2693–2705. https://doi.org/10.1007/s00894-011-0971-1
47. Lim J, Jeon S, Shin HY, Kim MJ, Seong YM, Lee WJ et al (2020) Case of the Index Patient Who Caused Tertiary Transmission of COVID-19 Infection in Korea: The Application of Lopinavir/Ritonavir for the Treatment of COVID-19 Infected Pneumonia Monitored by Quantitative RT-PCR. J Korean Med Sci 35(6):e79–e79. https://doi.org/10.3346/jkms.2020.35.e79
48. Mothay D, Ramesh KV (2020) Binding site analysis of potential protease inhibitors of COVID-19 using AutoDock. Virusdisease 31(2):194–199. https://doi.org/10.1007/s13337-020-00585-z
49. Günther S, Reinke PY, Fernández-García Y, Lieske J, Lane TJ, Ginn HM et al (2021) X-ray screening identifies active site and allosteric inhibitors of SARS-CoV-2 main protease. Science 372(6542):642–646. https://doi.org/10.1126/science.abf7945
50. Goyal B, Goyal D (2020) Targeting the dimerization of the main protease of coronaviruses: a potential broad-spectrum therapeutic strategy. ACS Comb Sci 22(6):297–305. https://doi.org/10.1021/acscombsci.0c00058
51. Wyatt PG, Woodhead AJ, Berdini V, Boulstridge JA, Carr MG, Cross DM et al (2008) Identification of N-(4-piperidinyl)-4-(2,6-dichlorobenzoylamino)-1 H-pyrazole-3-carboxamide (AT7519), a novel cyclin dependent kinase inhibitor using fragment-based X-ray crystallography and structure-based drug design. J Med Chem 51(16):4986–4999. https://doi.org/10.1021/jm800382h
52. Tomonari M, Chie T, Yong-Tae K, Hongfei W, Wataru N, Takaho T, Mikako S, Shigeyuki Y (2016) SARS-CoV 3CL protease cleaves its C-terminal autoprocessing site by novel subsite cooperativity. Proc Natl Acad Sci 113(46):12997–13002. https://doi.org/10.1073/pnas.1601327113