Flavonoid content of the Libyan *Onosma Cyrenaicum*: isolation, identification, electronic chemical reactivity, drug likeness, docking, and MD study

Ghazala Elkwafia, Najwa Mohamed, Fakhri Elabbar and Radwan Alnajjar

Department of Chemistry, Faculty of Science, University of Benghazi, Benghazi, Libya; Department of Medicinal Chemistry, Faculty of Pharmacy, University of Benghazi, Benghazi, Libya; Department of Chemistry, University of Cape Town, Rondebosch, South Africa

Communicated by Ramaswamy H. Sarma

**ABSTRACT**

In this work, an attempt to identify the flavonoid content of the Libyan *Onosma Cyrenaicum* led to the isolation of three flavonoids 7,8-dihydroxy-2-(4-hydroxyphenyl)-4H-chromen-4-one (GE-001), 5,7-dihydroxy-2-(3-hydroxy-4-methoxy phenyl)-4H-chromen-4-one (GE-002) and 5,7-dihydroxy-3-(4-hydroxy-phenyl)-4H-chromen-4-one (GE-003). The isolated compounds were characterized using 1H and 13C-NMR techniques. A further DFT study at B97-XD with 6-311+G** basis set in water was conducted to calculate the isolated compounds’ global and local reactivity descriptors and Fukui indices along with their antioxidant activity. The drug-likeness and bioactivity properties of the isolated compounds were estimated and discussed. Finally, GE-001, GE-002, and GE-003 were docked into HCV NS5B polymerase active site and this was followed by molecular dynamic simulation to certify the obtained docking result and to obtain the MM-GBSA free binding energy of the isolated compounds.

**ARTICLE HISTORY**

Received 29 December 2020
Accepted 24 February 2021

**KEYWORDS**

*Onosma Cyrenaicum*; Natural products; DFT; Hepatitis C; Docking; MD; MM-GBSA

1. Introduction

Hepatitis C virus (HCV) discovered in 1989; it is an affiliate of the *Hepacivirus* genus belongs *Flaviviridae* family with the *Hepacivirus* genus (Kazakov et al., 2015). HCV is the principal causative agent of hepatitis C. It classified as a positive-sense single-stranded RNA genome (Choo et al., 1989; Powdrill et al., 2010). According to the World Health Organization (WHO), about 180 million people worldwide are infected with HCV infection (Lemm et al., 2014).

Chronic HCV infection is the primary reason for chronic hepatitis, which leads to hepatocellular carcinoma at an occurrence of 4–5% and cirrhosis in ~20% of cases each year in infected patients (Maasoumy & Wedemeyer, 2012). Poor prognosis increasingly terminates into hepato-related severe illnesses such as cirrhosis, fibrosis, and hepatocellular carcinomas (Sofia et al., 2012). Unfortunately, there is no accessible prophylactic vaccine; for the past 15 years, the treatment of chronic hepatitis C has depended on market-available medicines such as ribavirin and pegylated interferon (IFN) (Pawlotsky, 2013). These drugs have unknown nonspecific mechanisms of action. The new therapies are in improvement that acts directly on the hepatitis C virus at different targets in the viral life cycle. Some of the new anti-HCV drugs, including other HCV replication inhibitors and protease inhibitors, have been reached clinical improvement (Pawlotsky, 2013).

Treatment with direct-acting antivirals (DAA) agents has intensely changed the results of hepatitis C. Indeed, the sustained viral response (SVR) rates have extended extraordinary levels (>95%) (Webster et al., 2015) without practical adverse actions. However, the cost is still one of the main barriers to reach hepatitis C eradication mostly in the middle- and low-income countries (Chhatwal et al., 2016).

The main challenge in developing novel HCV therapies has been the rise of resistance to direct-acting antiviral (DAA) drugs. HCV has a significant rate of replication, with ten visions created daily (Benhamou et al., 2009). The viral...
protein liable for replication is NS5B RNA-dependent RNA polymerase (RdRp), which lacks correcting ability and has a lot of error rate. Generally, mutated viruses have low replication suitability than wild-type virus and are currently in much lower amounts. However, when subjected to selection pressure such as the addition of a drug, the quantity of wild-type virus decreases, and the mutated virus gains replication fitness. Some of the subsequent mutations cause changes in the construction of the viral enzymes that DAA drugs act on, thus producing the virus resistant to these drugs (Sentandreu et al., 2008). Interestingly, Quercetagetin can inhibit HCV RNA-dependent RNA polymerase (RdRp) through blocking RNA binding to the viral polymerase enzyme, a mechanism related with broad genotypic activity against the construction of the viral enzymes that DAA drugs act on, thus producing the virus resistant to these drugs (Sentandreu et al., 2008). Interestingly, Quercetagetin can inhibit HCV RNA-dependent RNA polymerase (RdRp) through blocking RNA binding to the viral polymerase enzyme, a mechanism related with broad genotypic activity against numerous HCV strains and a lot barrier to resistance mechanisms (Ahmed-Belkacem et al., 2014; Powdrill et al., 2010).

These new treatments are highly costly with the possibility of developing multidrug resistance that will not be curable. However, the viral polymerase is the main component of any virus’s life cycle. They initiate essential roles that help the significant procession of viral proper transcription and replication of the genome. Subsequently, for this cause, viral polymerases have been used as a promising target for the design and discovery of potent antiviral treatments within productions concentrating on medicinal and pharmaceutical chemistry (Sesmero & Thorpe, 2015). The nonstructural 5B (NS5B) polymerase of HCV is positively recognized as a 55–66 kDa protein and positioned at the HCV’s C-terminus. It is an attractive target for antiviral intervention (Akher et al., 2019). The crystal structure of NS5B was successfully predicted in 1999 (Bressanelli et al. 1999) and discovered a right-handed arrangement accompanied via three intercalating subdomains known as the finger, palm and thumb domains (Sesmero & Thorpe, 2015). The entire NS5B structure consists of about 590 amino acids (Karam et al., 2014).

The Boraginaceae family contains a diversity of shrubs, trees and herbs totalling around 146 genera and 2000 classes found worldwide (Ahmed & Kordofani, 2012). About 53 kinds of the Boraginaceae family are grown in Libya (Feng et al., 2013), Some species such as Cynoglossum Clandestinum and Onosma Cyrenaicum (Figure 1) are endemic and found in the Wadi Al-Agar area (known as Al-Jabal Al-Akhdar). There is no specimen for these species have been collected in Libya (Alaib et al., 2017). The traditional use of Onosma Cyrenaicum as a stimulant in rheumatism, bladder pain, kidney irritation, palpitation of the heart, and the treatment of fevers, pain relief asthma and bronchitis. They have given the powdered leaves of Onosma Cyrenaicum to children for regulating urine output.

Worldwide several of Boraginaceae species have been used for wounds and burns healing in humans. This family’s wound healing capability is associated with their anti-inflammatory, antibacterial, antioxidant, and antiviral activities (Aliasl et al., 2015; Oza & Kulkarni, 2017).

Generally, flavonoids compose large polyphenolic compounds consisting of a benzo-γ-pyrene structure and are universally found in plants (Kumar & Pandey, 2013). Recent reports showed that flavonoids a group of secondary metabolites of phenolic nature and responsible for important protective pharmacological effects such as anti-atherogenic, anticaner, antiviral, and anti-inflammatory activities (Yang et al., 2011). Therefore, flavonoids are now considered a crucial benefit in various disciplines about medicinal, cosmological, nutraceutical, and pharmaceutical science (Akher et al., 2019).

This plant’s primary and secondary metabolites are flavonoids, alkaloids, phytosterols, terpenoids, and polyphenols (Dresler et al., 2017; Kocalf et al., 2012; Orhan, 2012). These compounds have antibiotic and antioxidant activities directly reliant on their phenolic compounds, such as phenolic acids and flavonoids (Ahmad et al., 2014; Moita et al., 2014). Besides, many stated biological effects, such as anti-inflammatory, antioxidant, anticaner, and antimutagenic activities, were credited to the presence of flavonoid and phenolic compounds are extensive in this family (Salem et al., 2014).

Quercetagetin (Que) is a natural flavonoid (Figure 2). It can be obtained from Citrus unshiu peel as citrus by-product observed as a strong protective effect on DNA damage caused via H2O2 (Yang et al., 2011). Additionally, Quercetagetin has several activities, such as antioxidant, antibacterial, and antifungal (Céspedes et al., 2006). Recently, Quercetagetin identified as the most potent non-nucleoside NS5B RdRp inhibitor among six flavonoid subfamilies experienced. Quercetagetin acts via inhibition of the RNA binding to the viral polymerase (Ahmed-Belkacem et al., 2014; Karam et al., 2014). Several other natural flavonoid compounds have been extracted, and some of them may indicate higher biological activity relative to the current inhibitors (Nyandoro et al., 2017; Toume et al., 2015; Xu et al., 2013).

Aiming to obtained new and competitive flavonoids compare to Quercetagetin, with ability to be developed into natural drug candidates the flavonoids contents of the Libyan Onosma Cyrenaicum were the objective of this study.
Successfully three compounds were isolated and identified (GE-001, GE-002, and GE-003); their electronic properties, physicochemical properties, drug-likeness and bioactivity properties were estimated. The isolated flavonoids were docked into the active site the nonstructural 5B polymerase (NS5B) of the hepatitis C virus (HCV) to get more details on isolated compounds’ ability to bind to NS5B protein active site. Finally, molecular dynamics simulations were conducted on the complexes obtained from the docking step for a 100 ns period, MD was used to monitor the stability of these complexes and the stability of the investigated compounds binding ability into the active site of the protein. Lastly the MM-GBSA binding energies for all complexes were reported.

2. Experimental and methods

The experimental procedure and the computational methods are described in details in the supporting information file.

3. Result and discussion

Libyan Onosma Cyrenaicum has been poorly explored, and no subspecies have been listed until today, also the natural organic compound contents of this plant still unrevealed. Attempting to focus on this plant’s content, we try to isolate the flavonoid contents hoping to get new flavonoids with potent biological activity and good physicochemical properties. This plant is seasonal, and it can only be collected during spring (March to May), this plant is also listed as distinct species. This plant is seasonal, and it can only be collected during spring (March to May), this plant is also listed as distinct species.

3.1. $^1$H and $^{13}$C-NMR analysis

The $^1$H-NMR spectrum of GE-001 is shown in Figure 4. The alpha proton H1 gives a singlet at $\delta$ 6.71 ppm with one hydrogen integration, protons H2 overlaps with protons H3 and H4 at $\delta$ 6.94 ppm given a doublet ($J$ = 8.68 Hz) with three hydrogen integration, this doublet is due to coupling with protons H5, H6, and H7, as it can be seen in Figure 4. H5 gives doublet at $\delta$ 7.39 ppm with $J$ = 8.63 Hz; also H6 and H7 give a doublet at $\delta$ 8.01 ppm with $J$ = 8.82 Hz, it looks like the doublet at 6.94 is actually two doublets overlapping at the same region. The phenolic hydroxides give a singlet signal at $\delta$ 9.38 ppm and $\delta$ 10.22 ppm with three protons’ total integration. The $^{13}$C-NMR spectra show 13 signals referring to the 15 carbons of the GE-001, the phenyl ring is symmetrical and hence it will only give four signals instead of six.

For GE-002, the 1H-NMR spectra showed three single signals at $\delta$ 12.13, 10.77, and 9.08 with one hydrogen integration for each representing the three phenolic OH respectively; the benzo group of the coumarin moiety hydrogens showed a multiple at $\delta$ 6.93 ppm with two hydrogen integration, while the phenyl ring gives two doublets the first at $\delta$ 6.87 ppm refreeing to proton H7 while the second doublet at $\delta$ 5.90 ppm refreeing to protons H5 and H7. The alphatic protons of the coumarin give three signals, doublet of a doublet at $\delta$ 5.43 with $J$ = 12.3 and 3.1 Hz and one proton integration presenting H2, this proton is coupled to protons H4 $\delta$ 3.19 (dd, $J$ = 17.1, 12.4 Hz) ppm and proton H1 $\delta$ 2.71(dd, $J$ = 17.1, 3.2 Hz); proton H4 is cis to H2 while H1 is trans is DFT calculations suggest. The methyl group appeared as single at $\delta$ 3.78 ppm. $^{13}$C spectra gives sixteen singles at $\delta$ 196.08, 166.57, 163.40, 162.72, 147.81, 146.40, 131.08, 117.57, 113.99, 111.91, 101.73, 95.73, 94.91, 78.14, 55.59, and 41.99 ppm; the $^1$H-NMR is presented in Figure 5.

GE-003 found to be isoflavone, and its $^1$H-NMR spectrum is presented in Figure 6. The spectra show three peaks at $\delta$ 12.96, 10.86, and 9.58 ppm for the hydroxy protons, the alpha proton (H7) gives a singlet at $\delta$ 8.31 ppm with one hydrogen integration, the phenyl group shows two doublets at $\delta$ 7.38 and 6.83 ppm respectively with $J$ = 8.7 Hz; finally, the benzo ring present two doublets with meta coupling $J$ = 2.1 Hz at 6.38 and 6.23 ppm respectively with one hydrogen each. The $^{13}$C-NMR shows peaks at $\delta$ 180.13, 164.19, 161.92, 157.50, 157.34, 153.85, 130.06, 122.21, 121.13, 114.98, 104.39, 98.88, and 93.57 presenting the fifteen carbons with four carbons for the phenyl moiety due to symmetry.

3.2. DFT calculations

3.2.1. Molecular geometry

Since these compounds have flexible hydroxyl groups, a conformational analysis for GE-001 at \(\omega\)B97XD/6-31+G**/H2O has been conducted and three possible conformations regarding the substituted hydroxy groups were founds, these conformers were further optimized at \(\omega\)B97XD/6-311+G**/H2O, the relative energy of these hydroxy group conformers found to be small (\(-1.62\) kcal/mol) and it could neglect, the relative
energies and the optimized structures of these conformations is reported in Figure 1S (supplementary material).

The critical parameters such as bond length and bond angle of GE-001, GE-002, and GE-003 were obtained at ωB97XD combined with 6-311++G(2d,p) basis set in water. Bonds lengths and angles are reported in Tables SI4 and SI5 respectively, the bond lengths and bond angles were within normal range comparing to a similar previous available crystal of flavones (Theodoro et al., 2008) (GE-001 family) and flavanones (Białońska et al., 2007a; 2007b) (GE-002 family), no previous crystal data for isoflavones family were available (GE-003 family). The Dihedral angle of the phenyl group in GE-001 found to 24.04°, while in GE-003 it was 49.01°, this dihedral angle effect on the solubility of these molecules, the larger the angle, the lower the crystal packing and the higher the solubility; the carbon atom that connects the phenyl group to the chromen moiety in GE-002 is saturated with sp3 hybridization, so the dihedral angle was 112.93°. The optimized structures of isolated compounds are shown in Figure 7.

3.2.2. Frontier molecular orbitals

To determine the electron distribution, the HOMOs and LUMOs of the GE-001, GE-002, and GE-003, which shown in Figure 8. The HOMO in GE-001 is concentrated over the benzo ring and the hydroxy groups of the chromen moiety, while in GE-002 the is localized on the phenyl ring and the methoxy group, in the case of GE-003 the HOMO is delocalized over the whole molecule. On the other side, the LUMO in the three compounds is mainly focused on the carbonyl group with extended conjugation of the electron density to the molecule’s chromen moiety. The HOMO and LUMO diagram presented in Figure 8, the positive and negative phases are in red and green colors, respectively.

3.2.3. Electronic properties

The electronic properties of the isolated compounds such as ionization potential (I), electron affinity (A), global hardness (η), electronegativity (χ), the electronic chemical potential (μ), electrophilicity index (α), nucleophilicity (N) index, and chemical softness (S) are listed in Table 1. The energy of HOMO is directly related to the ionization potential (I), while the energy of LUMO is associated with the electron affinity (A). The global hardness (η) resembles the energy gap between HOMO and LUMO. A molecule with a small energy gap has high chemical reactivity, low kinetic stability, and a soft molecule, while a hard molecule has a large energy gap (Zhan et al., 2003). The calculated results are reported in Table 1.

The global hardness (η) parameter is related to the energy gap (Eg = ELUMO − EHOMO) and defined as the measurement from the resistance of an atom or a group of atoms to charge transfer. The Eg is an important parameter for determination of the reactivity of the compounds. The electronic transport at the molecule with a low gap is easier. A molecule with a low Eg has a high chemical reactivity, low kinetic stability, and a soft molecule, whereas a hard molecule has a large energy gap (Zhan et al., 2003). The calculated results are reported in Table 1.

The global hardness (η) parameter is related to the energy gap (Eg = ELUMO − EHOMO) and defined as the measurement from the resistance of an atom or a group of atoms to charge transfer. The Eg is an important parameter for determination of the reactivity of the compounds. The electronic transport at the molecule with a low gap is easier. A molecule with a low Eg has a high chemical reactivity, low kinetic stability, and a soft molecule, whereas a hard molecule has a large energy gap (Zhan et al., 2003). The calculated results are reported in Table 1.

The global hardness (η) parameter is related to the energy gap (Eg = ELUMO − EHOMO) and defined as the measurement from the resistance of an atom or a group of atoms to charge transfer. The Eg is an important parameter for determination of the reactivity of the compounds. The electronic transport at the molecule with a low gap is easier. A molecule with a low Eg has a high chemical reactivity, low kinetic stability, and a soft molecule, whereas a hard molecule has a large energy gap (Zhan et al., 2003). The calculated results are reported in Table 1.
Figure 5. $^1$H-NMR spectra of GE-002 with labelled atoms.

Figure 6. $^1$H-NMR spectra of GE-003 with labelled atoms.
7.94 eV, 8.03 eV and 8.02 eV for GE-001, GE-002, and GE-003 respectively, indicating good stability. The electrophilicity index ($\omega$) shows that GE-001, GE-002, and GE-003 are good electrophiles ($\omega > 1.50$ eV) $\omega = 2.36$ eV, 2.09 eV, and 2.13 eV, respectively (Domingo et al., 2002). While the nucleophilic index (N) indicating that GE-001 is a moderate nucleophile with $N = 2.79$ eV (3.00 < $N$ < 2.00 eV), while GE-002 and GE-003 were poor nucleophile with $N = 0.19$ eV and 0.16 eV respectively (N < 2.00 eV) (Jaramillo et al., 2008).

3.2.4. Molecular electrostatic potential (MEP) analysis
Molecular electrostatic potential (MEP) maps identify the negative and positive sites of electrostatic potentials for electrophilic and nucleophilic reactions. The color of the surface represents the difference in electrostatic potential. Blue refers to the most positive region followed by green, while yellow, orange and red represent the negative regions with red being the most negative. The isolated compounds’ MEPs were computed at ωB97X/311++G**/H₂O level of theory (Figure 9). As shown in Figure 9, the carbonyl group of GE-001 has the highest electron density, and the hydroxyl group’s hydrogen has the lowest electron density the rest of the molecule is neutral. GE-002 and GE-003 carbonyl group involved in hydrogen bond and hence they are less electron denser, again the hydrogen of the hydroxyl groups is the most electron density deficient area.

3.2.5. Atomic charges
The Mulliken charges were calculated at ωB97X/311++G** for all non-H atoms are listed in Table 2. GE-001 shows the most negative charge on C7 the carbonyl carbon atom while the most positive where on carbon connected the phenyl group to the chromen moiety C9, all other heavy atoms show negative charges except the carbonyl neighbor C4 and C3. On the other hand, GE-002 most positive atom was C2 the carbonyl neighbor group, while all other heavy atoms were negatively charged with C1 being the most negative atom. GE-003 shows positive charges on carbonyl group neighbors and the carbon atom that connect the phenyl group to the chromen moiety; all other heavy atoms were

| Property | GE-001 | GE-002 | GE-003 |
|----------|--------|--------|--------|
| Energy (a.u) | $-953.69480$ | $-1069.44488$ | $-953.70141$ |
| Dipole moment | 7.94 | 5.41 | 6.08 |
| EHOMO (eV) | $-8.30$ | $-8.11$ | $-8.14$ |
| ELUMO (eV) | $-0.36$ | $-0.08$ | $-0.12$ |
| $E_g$ (eV) | 7.94 | 8.03 | 8.02 |
| $\chi$ (eV) | 8.30 | 8.11 | 8.14 |
| $\eta$ (eV) | 0.36 | 0.08 | 0.12 |
| $\mu$ (eV) | $-4.33$ | $-4.10$ | $-4.13$ |
| $\omega$ (eV) | 2.36 | 2.09 | 2.13 |
| $S$ (eV) | 0.25 | 0.25 | 0.25 |
| $N$ (eV) | 2.79 | 0.19 | 0.16 |

Figure 7. Optimized structure of the isolated compounds at ωB97X/311++G**.

Figure 8. Frontier molecular orbitals of the isolated compounds calculated at ωB97X/311++G**/H₂O.

Table 1. The calculated electronic properties of the isolated compounds GE-001, GE-002, and GE-003 at ωB97X/311++G**/H₂O.
negatively charged. All hydrogens atoms in GE-001, GE-002, and GE-003 were positively charged.

3.2.6. Fukui descriptor

The Fukui function describes the electron density after adding or removing electrons. It could estimate the position of the most electrophilic and nucleophilic sites on a molecule (Parr & Yang, 1984).

The Fukui equation for the electrophilic attack, \( f^- \), is given for when one electron is removed, while the Fukui for nucleophilic attack \( f^+ \) is given when one electron is added.

\[
    f^- = \rho(N) - \rho(N - 1) \\
    f^+ = \rho(N + 1) - \rho(N)
\]

The combination of Fukui functions will lead to formation of the Dual descriptor, the positive value represents the electrophilic while the negative value represents nucleophilic. The Dual descriptors can be obtained using the following equation:

\[
    f(r) = f^+ - f^-
\]

The Fukui function can be used to describe local chemical reactivity. This can even be done per atom by using the condensed Fukui function. Yang and Mortier (Yang & Mortier, 1986), proposed an approach to calculate Fukui functions based on the variation of the Mulliken gross charges, \( q(r) \), of an atom in a molecule. In a finite difference approximation, the condensed Fukui functions are given by,

\[
    f^- = q_N(N) - q_N(N - 1) \ldots \text{for electrophilic attack} \\
    f^+ = q_N(N + 1) - q_N(N) \ldots \text{for nucleophilic attack}
\]

The Fukui condensed functions were calculated for the isolated compounds using the UCA-FUKUI software (Sánchez-Márquez et al., 2014) and are reported in Table 3. The Fukui functions showed that the GE-001 will work as nucleophile using C6 and C8, while the electrophilic part of the molecule is C15 and C19. GE-002 on the other hand while use C9 and C17 to attack electrophile and C1 and C2 to accept nucleophilic attacks. Finally, the Fukui functions for GE-003 showed that C12, C13 will be used for nucleophilic substitution reactions and C3 and C10 for electrophilic reactions.

3.2.7. Antioxidant properties

The antioxidant properties of the isolated compounds were calculated using hydrogen atom transfer (HAT) approaches:

\[
    \text{ArOH} \rightarrow \text{ArO}^* + \text{H}^+
\]

Using the Enthalpies (H) at STP of the reactants and products obtained by Gaussian using the following expression:

\[
    H(T) = E_0 + ZPE + H_{trans} + H_{rot} + H_{vib} + RT
\]

Where \( E_0 \) and ZPE are the total energy at 0 K and the zero-point vibrational energy, respectively. \( H_{trans}, H_{rot} \) and \( H_{vib} \) are the translational, rotational, and vibrational contributions to the enthalpy.

The reaction enthalpies (bond dissociation enthalpy (BDE)) of the antioxidants in water were calculated using the following equation:

\[
    \text{BDE} = H(\text{ArO}^*) + H(\text{H}^*) - H(\text{ArOH})
\]

The obtained BDE values are reported in Table 4. Lower BDE value describes a higher antioxidant capacity. All O-H bonds dissociation enthalpy (BDE) is calculated in the aqueous phase at the wB97XD/6-311++G(d,p) level of theory. The obtained results are listed in Table 3, and the optimized structures presented in Figure 10. It is observed that BDE values vary from 76.56 to 107.76 kcal/mol, and the BDE are in the following order GE-001 < GE-002 < GE-003.
indicating that the antioxidant activity is in the following order GE-001 > GE-002 > GE-003. GE-001 showed the lowest BDE with H24 being the easiest to dissociate, while GE-002 came next with H33 BDE of 80.92 kcal/mol; finally, GE-003 has the highest BDE, and the lowest antioxidant activity with H28 requiring 83.20 kcal/mol to dissociate. In general, the dissociation of the phenyl group’s hydroxyl hydrogens is more favored with the para position being preferable.

Table 3. The Fukui indices of isolated compounds calculated at wB97XD/6-311++G(d,p).

| Compound | Atom | Fukui condensed function \( \tilde{f}^+ \) | Fukui condensed function \( \tilde{f}^- \) | Fukui condensed function \( \tilde{f}^\ast \) | Condensed dual descriptor |
|----------|------|-----------------|-----------------|-----------------|---------------------|
| GE-001   | C15  | 0.2143          | –               | 0.1352          | 0.1582              |
|          | C19  | 0.3904          | –               | 0.2060          | 0.369               |
|          | C6   | –               | 0.0432          | –               | –                   |
|          | C8   | –               | 0.1042          | –               | –                   |
| GE-002   | C1   | 0.3936          | –               | 0.2004          | –                   |
|          | C2   | 0.3956          | –               | 0.1984          | –                   |
|          | C9   | –               | 0.4268          | 0.2137          | 0.1341              |
|          | C17  | –               | 0.4574          | 0.2419          | 0.1449              |
| GE-003   | C3   | 0.2139          | –               | 0.1075          | 0.2127              |
|          | C10  | 0.3337          | –               | 0.1707          | –                   |
|          | C12  | 0.4199          | 0.2456          | 0.1341          | –                   |
|          | C13  | 0.3723          | 0.191           | 0.1341          | –                   |
|          | C11  | –               | –               | –               | 0.1341              |

Table 4. The O – H bond dissociation enthalpy (BDE) of the isolated compounds calculated at wB97XD/6-311++G(d,p) in water as a solvent.

| Molecule | Rad-1 (BDE in kcal/mol) | Rad-2 (BDE in kcal/mol) | Rad-3 (BDE in kcal/mol) |
|----------|-------------------------|-------------------------|-------------------------|
| GE-001   | 76.56 (H24)             | 83.13 (H25)             | 85.10 (H30)             |
| GE-002   | 106.34 (H31)            | 107.76 (H32)            | 80.92 (H33)             |
| GE-003   | 83.20 (H28)             | 95.34 (H29)             | 90.51 (H30)             |

Figure 10. Possible radicals for each compound obtained at oB97XD/311++G** level.
3.3. Evaluation of drug-likeness (drug ability)

Water solubility and lipophilicity are the main molecular properties in the absorption of a drug. Pharmacokinetic properties were chief approached academically in 1997 when Lipinski and colleagues published the Rule of Five (Ro5) on the study of properties of 2245 drugs of the World Drug Index (WDI) databank accepted for Phase II clinical trials (Lipinski et al., 1997). It used to calculate the oral bioavailability of a drug, this rule depends on physicochemical characteristics of the experienced molecules, involving: (i) clogP ≤ 5; (ii) molecular weight (MW) ≤ 500 g/mol; (iii) number of hydrogen bond acceptors (HBA) (sum of N and O atoms) ≤ 10; (iv) number of hydrogen bond donors (HBD) (sum of OH and NH groups) ≤ 5. other related criteria were added later (Veber et al., 2002); (v) number of rotatable bonds (nRotb) ≤ 10; (vi) polar surface area (PSA) < 140 Å².

The physicochemical properties like molecular weight (MW), molecular refractivity (MR), partition coefficient (log P<sub>o/w</sub>), polar surface area (PSA), number of hydrogen bond acceptors and donors counts, and the number of rotatable bonds in a molecular count of specific atom types of the isolated compounds were estimated using SwissADME to evaluate the drug-likeness of isolated compounds and were reported in Table 5.

The degree of oral bioavailability of molecules is an important balance between molecular properties that affect the pharmacodynamics and pharmacokinetics of compounds which affects their absorption, distribution, metabolism, and excretion (ADME) drug-likeness within the human body (Buchwald & Bodor, 2002).

The molecular weight of all compounds (GE-001, GE-002, and GE-003) was less than 500 (270.24, 302.28, and 270.24), and lower than Que, which make it easy to cross membranes, transported, absorbed and diffused compared to heavy molecules (Waring, 2009). The isolated compounds showed enhancement in tPSA compared to Quercetagetin, Table 5, lower tPSA indicating longer half-time inside the body (Ertl et al., 2000). Since lipophilicity is hypothetically associated with toxicity, clogP considers a crucial factor in drug properties. Low clogP 0–3 indicates an equilibrium between absorption, excretion, and toxicity, and the isolated flavonoids showed balanced clogP (Arnott & Planey, 2012). A recent study demonstrates that compounds that exhibit a clogP between 1–3 showed to be optimal for reaching suitable physicochemical characteristics to ensure whole drug success (Arnott & Planey, 2012). The clogP values for isolated compounds and Quercetagetin were found to be 1.94, 1.91, 2.04, and 0.92, Table 5, which obey Lipinski’s rule of five, so all compounds have an adequate balance between lipophilicity and hydrophilicity; which observe that the compounds will have a good permeability across cell membranes (Ritchie et al., 2013). As reported from Table 5, Quercetagetin is more water-soluble (−3.96) than other isolated compounds due to its large number of hydroxyl groups compared to isolated compounds. There is a small difference in water solubility between isolated flavonoid compound GE-001 and GE-003 (−4.04, −4.23), respectively, due to the same number of hydroxyl groups in a different position. Finally, GE-002 has the lowest water solubility (−4.27) since it has an ether group instead of the hydroxyl group, the ether oxygen is a hydrogen-bond acceptor, while the hydroxyl group is hydrogen bond donor and acceptor.

Table 5. physicochemical properties of isolated compounds.

| Molecule  | MW    | MR   | tPSA  | clogP  | logS  | HBA | HBD | nRotb | Bioavailability score |
|-----------|-------|------|-------|--------|------|-----|-----|-------|----------------------|
| GE-001    | 270.24| 73.99| 90.90 | 1.94   | −4.04| 5   | 3   | 1     | 0.55                 |
| GE-002    | 302.28| 78.06| 96.22 | 1.91   | −4.27| 6   | 3   | 2     | 0.55                 |
| GE-003    | 270.24| 73.99| 90.90 | 2.04   | −4.23| 5   | 3   | 1     | 0.55                 |
| Que       | 318.24| 80.06| 151.59| 0.92   | −3.96| 8   | 6   | 1     | 0.55                 |

hAtoms: heavy atoms; HBA: number of hydrogen bond acceptor; HBD: number of hydrogen bond donor; nRotb: number of rotatable bonds.

3.4. Pharmacokinetics

As it was concluded from the previous results, isolated flavonoids have physicochemical properties within a suitable range. Therefore, these properties should be occupied into reflection as an indicator for further estimation investigates against different targets involving GI Absorption, BBB
permeant, P-gp substrate, CYP1A2 inhibitor, CYP2C19 inhibitor, CYP2C9 inhibitor, CYP2D6 inhibitor, CYP3A4 inhibitor, Log Kp; consequently, by using SwissADME, the bioactivity of all molecules were corroborated and described in Table 6. The evaluations for passive human gastrointestinal absorption (GI Absorption) of the isolated compounds, observed that all isolated compounds have higher GI absorption than Quercetagetin. Moreover, the blood-brain barrier (BBB permeant) gives to brain homeostasis via protective the brain from possibly harmful endogenous and exogenous materials, from (Table 6) presented that all molecules cannot cross BBB, that all isolated compounds have higher GI absorption than Quercetagetin. Moreover, the blood-brain barrier (BBB permeant) gives to brain homeostasis via protective the brain from possibly harmful endogenous and exogenous materials.

### Table 6. Pharmacokinetics for Quercetagetin and isolated flavonoid compounds.

| Compounds | GI Absorption | BBB permeant | P-gp substrate | CYP1A2 inhibitor | CYP2C19 inhibitor | CYP2C9 inhibitor | CYP2D6 inhibitor | CYP3A4 inhibitor | Log Kp |
|-----------|---------------|--------------|----------------|------------------|------------------|------------------|------------------|------------------|--------|
| GE-001    | High          | No           | No             | Yes              | No               | No               | Yes              | Yes              | –6.18  |
| GE-002    | High          | No           | Yes            | Yes              | No               | No               | Yes              | Yes              | –6.30  |
| GE-003    | High          | No           | Yes            | No               | No               | Yes              | Yes              | Yes              | –6.05  |
| Que       | Low           | No           | No             | Yes              | No               | No               | Yes              | Yes              | –7.40  |

3.5. Molecular docking

The active site of HCV NS5B polymerase is located inside the protein. It is situated between two beta-sheets and a loop. The active site is not directly exposed on the surface of the protein, Figure 11. The active site consists of electronically charged amino acids (His95, Arg559, Glu446, and Asp444), polar amino acids (Pro93, Pro94, Ser96, Gly283, Pro405, Gly449, Gly557, Gly558, and Cys451), and Lipophilic amino acids (Ala97, Phe162, Ile405, and Ile560). All isolated compounds were docked into the protein’s active site, and the obtained results are presented in Table 7.

As shown in Table 7, Que still holds the best result with a –6.05 kcal/mol docking score, the isolated compounds were in the following order GE-002 > GE-001 > GE-003. The keto group of the chromanone moiety in GE-001 form an H-bond with Asp559 at 2.17 Å. GE-002 and GE-003, on the other hand, used the hydroxyl groups on the benzo moiety of the chromanone, GE-002 used the 5-hydroxy group to form an H-bond with Ser96 side chain at 2.30 Å, it also from a C-H to π interaction with Gly557, while GE-003 used the 7-hydroxy group to interact via H-bond with Gly557 at 2.17 Å, and forms two π – C-H interactions with Asp559 and His95. Finally, Que creates two H-bonds with Arg158 and Asp318, at 2.19 and 2.29 Å, respectively, using the keto group and 6-hydroxy groups of the chromanone moiety. In general, the phenyl group of all compounds was not involved in any interactions with the protein residues due to the pocket’s polar nature. The poses with the best score are presented in Figure 12.

3.6. Molecular dynamics simulations

#### 3.6.1. Protein and ligand RMSD analysis

The conformational stability of the protein was mentioned through the RMSD of the Cα atoms of the protein with respect to their initial structure. The obtained RMSD value was plotted as a function of the simulation time. As it can be seen from Figure 13, the protein RMSD was within acceptable range < 3.00 Å; most complexes stabilized at around 60 ns of the simulation time to fluctuate at around 1.50 – 2.00 Å from its initial atom position. The RMSD for all compounds is shown in Figure 13.

The active site RMSD indicates how stable the ligand is with respect to the protein and its binding pocket (ligand fit on protein); it is plotted as a function of time in a...
nanosecond and presented in Figure 14; as it can be observed the GE-001 is the most stable compounds within the active site, it moves only around 1.50 Å of its initial position, GE-003 was moved \( \sim 6.00 \) Å before reach equilibrium at around 10 ns of the simulation time. Followed was GE-003 which require 20 ns to achieve stability in the active site, GE-002 fluctuation till 30 ns before reaching stability, GE-002 is fluctuated regularly due to the saturated carbon atom which increases the volume of GE-002 and reduces its planarity. Que was fluctuated until 50 ns seconds before becoming stable within the active site.

The phenyl group rotation was also monitored during the simulation time, the average dihedral angle for GE-001, GE-002, GE-003, and Que were around \( \sim 43^\circ, 26^\circ, 119^\circ \) and \( 142^\circ \), respectively, Figure 15 and Figure SI2 (supplementary material).

The finally equilibrated RMSD (FE-RMSD) of the protein Cα atoms and the ligand (ligand atoms as reference) and ligand-protein active site for each complex were defined as their average values in the last ten ns and reported in Table 8. For all the complexes, FE-RMSD of protein and ligand are \( \leq 3 \) Å, validating the MD simulations. While the active site FE-RMSD is in the following order GE-001 < Que < GE-003 < GE-002.

Since docking is not reliable as it misses the protein movement, molecular dynamics was used to study the interactions that ligand forms with the active site residuals, Figure16 represent these interactions over time, only interactions that occurred during 33% of the simulation time were considered. GE-001 was able to form four H-bonds whit Ala96, Val161, Ser282, and Asp559 during 90%, 87%, 82%, and 98% of the time, respectively. Some lipophilic interactions were also formed with Lys141 (37%), Ile160 (43%), and Phe162 (40%). GE-002 was able to form three H-bonds with the following residuals Glu17 (62%), Ala97 (43%), and Asp(33%), and most of its lipophilic interactions were less the 30% of the simulation time except for Phe162 (37%). Also, another H-bond was formed through water bridge with Gly557 during 32% of the time. GE-003 formed only one H-bond with Ser225 during 71% of the simulation time, two lipophilic interactions with Phe160 and Phe162 during 55% and 86% of the time, respectively, and one hydrogen bond through water bridge with Lys141 during 42% of the time. Finally, Que formed four H-bonds with the following residuals Ala97 (44%), Gly557 (53%), and Asp559 (79%), one lipophilic interaction with Phe162 (66%), and three H-bonds through water bridge with Ala140 (44%), Ser142(55%), and Gly557(38%).

### 3.6.2. MM-GBSA calculations

The effective binding free energies (\( \Delta G_{\text{eff}} \)) of the complexes between the GE-001, GE-002, GE-003, and Que and NS5B were computed considering the solvation free energy contributions to binding using the MM-GBSA approach. For this, structural ensembles for each complex were extracted every two ns from the 100 ns simulations. The relevant module in Schrodinger Suite was used, i.e. the thermal_mmgbsa.py script that takes individual trajectory snapshots and calculates \( \Delta G_{\text{eff}} \) and its energetic contributions

The binding energy is calculated according to the equation:

\[
\Delta G_{\text{bind}} = E_{\text{minimized complex}} - E_{\text{minimized ligand}} - E_{\text{minimized receptor}}
\]

The obtained MM-GBSA free energy of binding was in the following order GE-001 > Que > GE-003 > GE-002, as it can be seen in Table 9. GE-001 also showed better coulomb energy among the isolated compound and relative to the Que coulomb energy and showed the best van der Waal’s interaction energy among all compounds.
4. Conclusion

In this work, the flavonoid contents extraction of the Libyan Onosma Cyrenaicum led to identifying of three compounds GE-001, GE-002, and GE-003. These compounds were identified using NMR spectroscopy. The isolated compounds were subject to DFT calculations at \( \omega B97-XD \) with 6-311++G** level of theory to obtain their electronic and antioxidant properties, and the three compounds showed good antioxidant properties. The drug-likeness and...
Figure 13. Plots of RMSD for Cx atoms (Å) with respect to the initial structure vs simulation time (ns) for all the complexes.

Figure 14. RMSD of the atomic positions of the ligands fitting the receptor’s active site of the 100 ns molecular dynamics simulations for all the complexes.

Figure 15. Overlay of ligands trajectories during the simulation time.
pharmacokinetic properties were evaluated, and the obtained properties found to be promised, all compounds obey Lipinski’s rule of five, which Que failed to obey.

A docking study against HCV NS5B polymerase showed that Que has a higher score than the isolated compounds.

Figure 16. Receptor–ligand histogram interaction plot of GE-001, GE-002, GE-003, and Que conformation inside NS5B for 100 ns of MD simulations. Hydrogen bonding interaction bar is depicted in light blue, van der Waals in yellow, and water bridges in blue.

Table 8. The finally equilibrated values of RMSD (FE-RMSD) for each complex.

| Regions       | NS5B-GE-001 | NS5B-GE-002 | NS5B-GE-003 | NS5B-Que |
|---------------|-------------|-------------|-------------|----------|
| Protein-FE-RMSD (Å) | 2.41        | 2.15        | 2.38        | 1.90     |
| Ligand-FE-RMSD (Å)  | 0.25        | 1.36        | 2.14        | 0.65     |
| Active site – FE-RMSD (Å) | 1.21        | 8.11        | 5.50        | 4.41     |
However, a further investigation using molecular dynamic simulations showed that GE-001 was able to form more H-bonds than GE-002 and maintain these bonds more than 30% of the simulation time. Finally, MM-GBSA free energy of binding showed that GE-001 is more potent than GE-002 and have higher binding energy. In conclusion, these flavonoids showed promising activity against HCV, and could be considered for in vitro and in vivo studies.

**References**

Ahmad, S., Ahmad, S., Bibi, A., Ishaq, M. S., Afridi, M. S., Kanwal, F., Zakir, M., & Fatima, F. (2014). Phytochemical analysis, antioxidative activity, fatty acids composition, and functional group analysis of Heliotropium bacciferum. *The Scientific World Journal*, 2014, 1–8. https://doi.org/10.1155/2014/829076

Ahmed, H. O., & Kordofani, M. A. (2012). Leaf and stem anatomy of five species from the genus Heliotropium L.(Boraginaceae) in Sudan. *Journal of Chemical and Pharmaceutical Research*, 4(10), 4575–4581.

Ahmed-Belkacem, A., Guichou, J.-F., Brillet, R., Ahnou, N., Hernandez, E., Bressanelli, S., Tomei, L., Roussel, A., Incitti, I., Vitale, R. L., Mathieu, M., De Francesco, R., & Rey, F. A. (1999). Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Proceedings of the National Academy of Sciences of the United States of America*, 96(23), 13034–13039. https://doi.org/10.1073/pnas.96.23.13034

Buchwald, P., & Bodor, N. (2002). Computer-aided drug design: The role of quantitative structure–property, structure–activity and structure–metabolism relationships (QSPR, QSAR, QSMR). *Drugs of the Future*, 27(6), 577–588. https://doi.org/10.1358/dfb.2002.027.06.856934

Céspedes, C. L., Avila, J. G., Martínez, A., Serrato, B., Calderón-Mugica, J. C., & Salgado-Garciglia, R. (2006). Antifungal and antibacterial activities of Mexican tarragon (Tagetes lucida). *Journal of Agricultural and Food Chemistry*, 54(10), 3521–3527. https://doi.org/10.1021/jf053071w

Chhatwal, J., He, T., & Lopez-Olivo, M. A. (2016). Systematic review of modelling approaches for the cost effectiveness of Hepatitis C treatment with direct-acting antivirals. *Pharmacoconomics*, 34(5), 551–567. https://doi.org/10.1007/s40273-015-0373-9

Choo, Q.-L., Kuo, G., Weiner, A. J., Overby, L. R., Bradley, D. W., & Houghton, M. (1989). Isolation of a CDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science (New York, N.Y.*)*, 244(4902), 359–362. https://doi.org/10.1126/science.2523562

Clamp, M., Cohn, A. J., Pachter, L., & Gingeras, T. R. (2003). An introduction to the source code of the BLAST programs. *Theoretical Computer Science*, 306(1), 7–21. https://doi.org/10.1016/S0304-3975(02)00538-4

Drews, G. (2007). *Biotechnological Production of Plant Secondary Metabolites*. John Wiley & Sons.

Ertl, P., Rohde, B., & Selzer, P. (2000). Fast calculation of molecular polar surface area as a sum of fragment-based contributions and its application to the prediction of drug transport properties. *Journal of Medicinal Chemistry*, 43(20), 3714–3717. https://doi.org/10.1021/jm000942e

Feng, Y., Lei, J.-Q., Xu, X.-W., & Pan, B.-R. (2013). Composition and characteristics of Libyan flora. *Archives of Biological Sciences*, 65(2), 651–657. https://doi.org/10.2298/ABS1302651Y

Jaramillo, P., Domingo, L. R., Chamorro, E., & Pérez, P. (2008). A further exploration of a nucleicphile index based on the gas-phase ionization potentials. *Journal of Molecular Structure: THEOCHEM*, 865(1-3), 68–72. https://doi.org/10.1016/j.theochem.2008.06.022

Karam, P., Powdrill, M. H., Liu, H.-W., Vasquez, C., Mah, W., Bernatchez, J., Götze, M., & Cosa, G. (2014). Dynamics of hepatitis C virus (HCV) RNA-dependent RNA polymerase NS5B in complex with RNA. *J Biol Chem*, 289(20), 14399–14411. https://doi.org/10.1074/jbc.M113.529743

Kazakov, T., Yang, F., Ramanathan, H. N., Kohliway, A., Diamond, M. S., & Lindenbach, B. D. (2015). Hepatitis C virus RNA replication depends on specific cis- and trans-acting activities of viral nonstructural proteins. *PLoS Pathogens*, 11(4), e1004817 https://doi.org/10.1371/journal.ppat.1004817

Kocak, U., Citilgecen, H., & Reheman, N. (2012). Progress in biotechnological applications of diverse species in Boraginaceae Juss. *Biotechnological Production of Plant Secondary Metabolites*, Chapter 12, 200.

Kumar, S., & Pandey, A. K. (2013). Chemistry and biological activities of flavonoids: An overview. *The Scientific World Journal*, 2013, 1–16. https://doi.org/10.1155/2013/162750

Lemm, R. A., Liu, M., Gentles, R. G., Ding, M., Voss, S., Pelosi, L. A., Wang, Y.-K., Rigat, K. L., Mosure, K. W., Bender, J. A., Knipe, J. O., Colombo, R., Meanwell, N. A., Kadow, J. F., Santone, K. S., Roberts, S. B., & Gao, M. (2014). Preclinical characterization of BMS-791325, an allosteric inhibitor of hepatitis C virus NS5B polymerase. *Antimicrobial Agents and Chemistry*, 58(6), 3485–3495. https://doi.org/10.1128/AAC.02495-13

Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility

**Table 9. Calculated MM-GBSA binding free-energy in kcal/mol of the Studied Complexes.**

| Complex   | \( \Delta G_{\text{Bind}} \) | \( \Delta G_{\text{Coulomb}} \) | \( \Delta G_{\text{H-bond}} \) | \( \Delta G_{\text{Lipo}} \) | \( \Delta G_{\text{GB/soln}} \) | \( \Delta G_{\text{GB/DW}} \) |
|-----------|----------------|----------------|----------------|----------------|----------------|----------------|
| NS5B-GE-001 | -49.10 | -22.12 | -2.64 | -12.01 | 18.00 | -30.32 |
| NS5B-GE-002 | -35.36 | -13.74 | -1.32 | -8.46 | 16.78 | -28.71 |
| NS5B-GE-003 | -36.26 | -14.02 | -0.89 | -11.03 | 16.76 | -27.28 |
| NS5B-Que | -44.79 | -23.93 | -2.89 | -13.22 | 24.18 | -29.38 |

Whereas \( \Delta G_{\text{Bind}} \): free energy of binding; \( \Delta G_{\text{Coulomb}} \): coulomb energy of the complex; \( \Delta G_{\text{H-bond}} \): Hydrogen – bonding correction; \( \Delta G_{\text{Lipo}} \): Lipophilic energy; \( \Delta G_{\text{GB/soln}} \): Generalized Born electrostatic solvation energy.

Białońska, A., Ciunik, Z., Koztrzezwa-Suslow, E., & Dmochowska-Gładysz, J. (2007b). Disordered 6-hydroxyflavanone. *Acta Crystallographica Section E Structure Reports Online*, 63(2), o432–o433. https://doi.org/10.1107/S1600536806054687

Bresanelli, S., Tomei, L., Roussel, A., Incitti, I., Vitale, R. L., Mathieu, M., De Francesco, R., & Rey, F. A. (1999). Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. *Proceedings of the National Academy of Sciences of the United States of America*, 96(23), 13034–13039. https://doi.org/10.1073/pnas.96.23.13034

ORCID

Radwan Alnajjar http://orcid.org/0000-0003-1869-9596
