



This is a repository copy of *Interplay between  $n \rightarrow \pi^*$  interaction and hydrogen bond in an analgesic drug salicin.*

White Rose Research Online URL for this paper:  
<http://eprints.whiterose.ac.uk/132349/>

Version: Accepted Version

---

**Article:**

Singh, S.K., Joshi, P.R., Shaw, R. et al. (2 more authors) (2018) Interplay between  $n \rightarrow \pi^*$  interaction and hydrogen bond in an analgesic drug salicin. *Physical Chemistry Chemical Physics*, 20 (27). pp. 18361-18373. ISSN 1463-9076

<https://doi.org/10.1039/C8CP00655E>

---

© 2018, Royal Society of Chemistry. This is an author produced version of a paper subsequently published in *Physical Chemistry Chemical Physics*. Uploaded in accordance with the publisher's self-archiving policy.

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# **Interplay between hydrogen bonding and $n \rightarrow \pi^*$ interaction in an analgesic drug salicin**

Santosh K. Singh<sup>†</sup>, Prasad Ramesh Joshi<sup>†</sup>, Robert A. Shaw<sup>‡</sup>, J. Grant Hill<sup>\*‡</sup>,  
and Alope Das<sup>\*†</sup>

<sup>†</sup>Department of Chemistry, Indian Institute of Science Education and Research, Dr. Homi Bhabha Road, Pashan, Pune-411008, Maharashtra, India

<sup>‡</sup> Department of Chemistry, University of Sheffield, Sheffield S3 7HF, UK

## Abstract

The competition and cooperation between weak intermolecular interactions are important in determining the conformational preferences of molecules. Understanding the relative strengths of these effects in the context of potential drug candidates is therefore essential. We use a combination of gas-phase spectroscopy and quantum-chemical calculations to elucidate the nature of such interactions for the analgesic salicin [2-(hydroxymethyl)phenyl  $\beta$ -D-glucopyranoside], an analog of aspirin found in willow bark. Of several possible conformers, only three are observed experimentally, and these are found to correspond with the three lowest energy conformers obtained from density functional theory calculations and simulated Franck-Condon spectra. Natural bond orbital analyses show that these are characterized by a subtle interplay between weak  $n \rightarrow \pi^*$  interaction and conventional strong hydrogen bond, with additional insights into this interaction provided by analysis of quantum theory of atoms in molecules and symmetry-adapted perturbation theory calculations. In contrast, the higher energy conformers, which are not observed experimentally, are mostly stabilized by the hydrogen bond with negligible contribution of  $n \rightarrow \pi^*$  interaction. The  $n \rightarrow \pi^*$  interaction results in a preference for the benzylic alcohol group to adopt a gauche conformation, a characteristic also found when salicin is bound to the  $\beta$ -glucosidase enzyme. As such, understanding the interplay between these weak interactions has significance in the rationalization of protein structures.

## Introduction

The delicate interplay between multiple weak non-covalent interactions is vital in describing the structures and functions of biomolecules.<sup>1, 2</sup> Understanding the subtle balance among these weak interactions is a formidable but important task. The  $n \rightarrow \pi^*$  interaction, which is analogous to the hydrogen bonding interaction in terms of electron delocalization, is abundant in proteins, nucleic acids, neurotransmitters, the popular drug aspirin, biologically relevant small molecules, supramolecules etc.<sup>3-29</sup> However, unlike the hydrogen bonding interaction, this  $n \rightarrow \pi^*$  noncovalent interaction has not been explored by the scientific community until recently due to its weak and counterintuitive nature. The  $n \rightarrow \pi^*$  interaction is the delocalization of lone pair electrons (n) into the  $\pi^*$  orbital of a carbonyl group or an aryl group and thus it is counterintuitive due to the requirement of close proximity of an electronegative atom and a  $\pi$ -electron cloud.<sup>3, 4</sup> This non-covalent interaction is roughly 0.2–1 kcal/mol in magnitude and follows the Burgi-Dunitz trajectory for nucleophilic addition towards an electrophile.<sup>3, 4</sup> Previous symmetry-adapted perturbation theory (SAPT) decompositions of the interaction energy of  $n \rightarrow \pi^*$  interactions have focused on prototypical intermolecular complexes, such as benzene $\cdots$ dimethylether or benzene $\cdots$ trimethylammonia.<sup>30, 31</sup> In such cases it has been demonstrated that while the interaction relies on a mixture of components, there tends to be a relatively large dispersion term and a repulsive electrostatic contribution.

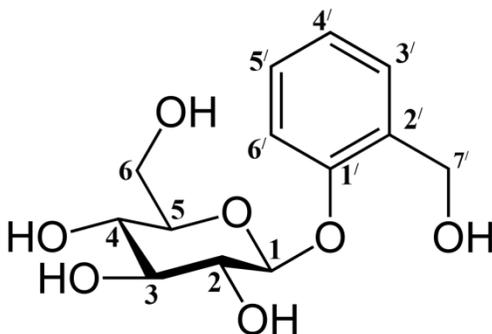
The  $n \rightarrow \pi^*$  interaction has been identified mostly through the analysis of X-ray crystal structures in the Cambridge Structural Database (CSD) and protein data bank (PDB) as well as NMR spectroscopy, X-ray diffraction and quantum chemical calculations of small molecules, such as peptides, peptoids, amino acids, amides and others.<sup>5, 7, 14, 16, 17, 24</sup> The first gas-phase IR spectroscopic evidence for an  $n \rightarrow \pi^*$  interaction was reported very recently by Das and co-workers.<sup>32</sup> It has also been shown recently that an  $n \rightarrow \pi^*$  interaction between two neighboring carbonyl groups can be reciprocal in nature, i.e. each carbonyl acts as both donor and acceptor in concurrent  $n \rightarrow \pi^*$  interactions.<sup>33</sup>

Interestingly, as both  $n \rightarrow \pi^*$  interactions and hydrogen bonding entail lone pair electrons (n) on a carbonyl oxygen atom, these two noncovalent interactions are interrelated. That is, the occurrence of one of the interactions can create the possibility of the existence of the other

interaction. Raines and co-workers have stated that there is a close connection between the hydrogen bond and an  $n \rightarrow \pi^*$  interaction,<sup>5</sup> and interplay or cooperativity between these two interactions is inevitable in the structures of biomolecules. It has been found that since there are two lone pairs (s- and p-types) on the oxygen atom, there are two main possibilities. The first is that one of the lone pairs participates in the hydrogen bonding while the other lone pair on the same atom takes part in the  $n \rightarrow \pi^*$  interaction, alternatively the same lone pair on the oxygen atom is simultaneously shared between both the hydrogen bonding and  $n \rightarrow \pi^*$  interactions.<sup>5, 6</sup> Here the presence of the  $n \rightarrow \pi^*$  interaction weakens the hydrogen bond and vice versa. The interesting point is that the hydrogen bond alone could be stronger than the combination of the compromised hydrogen bond and the  $n \rightarrow \pi^*$  interaction. However, the latter arrangement may be preferable from an optimal structure point of view.<sup>11</sup> Spectroscopic studies on how the interplay between the hydrogen bond and the  $n \rightarrow \pi^*$  interaction can dictate the structures of molecular systems are scarce in the literature.

In this work, the structure of an analgesic drug, salicin, has been studied and it is shown to be stabilized by  $n \rightarrow \pi^*$  interactions in addition to hydrogen bonds. Pharmacological activities of salicin are similar to those of the popular drug aspirin.<sup>34-39</sup> It has been reported by Raines and coworkers that  $n \rightarrow \pi^*$  interaction plays a significant role in controlling the preferred structure and biological activity of aspirin.<sup>21</sup> Generally, the clinical efficacy of a drug depends upon its interaction with biomolecules, which further depends upon the structure of the drug molecule and its possible conformations. In fact, medicinal properties of salicin are manifold. Salicin is the major chemical constituent of white willow bark (*salix alba*), also known as “nature’s aspirin”, which has been used as a medicine for the treatment of fever, pain, and inflammation in the body for thousands of years.<sup>34-38</sup> Salicin is a glycoside and its chemical structure (Figure 1) consists of a sugar molecule ( $\beta$ -D-glucopyranoside) and a benzyl alcohol moiety, which are linked through the anomeric oxygen atom of the sugar molecule. Salicin acts as a prodrug which is metabolized in our body to form the drug salicylic acid.<sup>34, 35, 37-39</sup> Upon ingestion, salicin undergoes hydrolysis to form salicyl alcohol (saligenin), which is further oxidized to generate the salicylic acid that acts as an antipyretic, analgesic, and anti-inflammatory agent in the body.<sup>37, 39</sup> Historically, the discovery of aspirin (acetyl salicylic acid) was triggered by the extraction of salicin from willow bark and subsequent synthesis of salicylic acid from salicin.<sup>37</sup> Recently it has been reported that salicin also shows antitumor properties by inhibiting angiogenesis, a process which supplies

oxygen and nutrients to tumor cells.<sup>40, 41</sup> Several in vitro and in vivo clinical trials are being carried out to discover and understand its full range of clinical properties.<sup>41-44</sup>



**Figure 1.** Skeletal structure of salicin showing atom-numbering scheme

Herein, we have explored the conformational preferences of salicin in isolated gas phase using mass-selected Resonance Two-Photon Ionization (R2PI), IR-UV and UV-UV double resonance spectroscopic techniques in combination with density functional theory calculations and calculated Franck-Condon spectra. Three low energy conformers of salicin have been observed in the experiment. These conformers show a chain of weak hydrogen bonding interactions between the equatorial hydroxyl groups of the sugar unit while the  $-\text{CH}_2\text{OH}$  group of the benzyl alcohol moiety gives rise to an additional strong  $\text{O-H}\cdots\text{O}$  hydrogen bonding interaction with one of the  $-\text{OH}$  groups of the sugar moiety. In addition to this, the oxygen atom of the  $-\text{CH}_2\text{OH}$  group of the benzyl alcohol moiety takes part in an  $n\rightarrow\pi^*$  interaction with the phenyl ring. It is noticed that a few higher energy conformers of salicin, possessing stronger hydrogen bonds but negligible  $n\rightarrow\pi^*$  interactions, are not observed in the experiment. Gas phase spectroscopic studies of salicyl alcohol (saligenin), which is the hydrolyzed product of salicin, have been reported in the literature.<sup>45, 46</sup> However, the structure of saligenin is much simpler than that of salicin.

The present research demonstrates that the conformational preferences of salicin are governed by an interplay between weak  $n\rightarrow\pi^*$  interaction and strong hydrogen bonding

interactions in the molecule. A close analog of salicin is phenyl  $\beta$ -D-glucopyranoside, which has been studied by Simons and co-workers as well as other groups using gas phase laser spectroscopy.<sup>47-49</sup> However, phenyl  $\beta$ -D-glucopyranoside lacks the strong intramolecular hydrogen bonding interaction as well as an  $n \rightarrow \pi^*$  interaction due to absence of the  $-\text{CH}_2\text{OH}$  group at 2' position of the aromatic ring and these missing interactions form the basis of the major focus of the current work.

## 2. Experimental

### 2.1. Experimental Methods

The experimental setup used in this work is similar to what has been described in previous publications.<sup>50-53</sup> However, the time-of-flight mass spectrometer used here was modified to incorporate a home-built laser desorption source. Salicin (vapor pressure =  $4 \times 10^{-13}$  mmHg at 25 °C) was brought into the gas phase using a laser desorption technique.<sup>54, 55</sup> A sample pellet (12 mm diameter and 2 mm thickness) was made by pressing a 2:1 mixture (weight ratio) of salicin (Sigma Aldrich, 99% Purity) and graphite powder (Sigma Aldrich, size ~20 micron) in a hydraulic press under 2-3 tons of pressure. The pellet was placed in a sample holder attached with an XYZ manipulator coupled with a motorized translational assembly (Fourvac Technology, Pune) for translation of the sample pellet along the Z-axis. Second harmonic output (532 nm) of a Nd:YAG laser (Continuum, Minilite-I, 10 Hz, 10 nanosecond) was used for desorption of the sample from the surface of the pellet. The 532 nm laser beam of about 500  $\mu\text{J}$  pulse energy was fed into an optical fiber (400  $\mu\text{m}$  core diameter, 4 m length) through an optical coupler and focused on the sample. The distance between the edge of the pellet and the center of the orifice of the pulsed valve was maintained at about 1 mm while the distance between the surface of the pellet and the axis of the molecular beam was kept at about 2 mm. The sample pellet was translated back and forth to ensure exposure of a fresh spot of the sample surface by the laser. The desorbed molecules were entrained in a supersonic expansion of Ar gas (~ 5 bar) through a pulsed nozzle of 500  $\mu\text{m}$  diameter (General valve, series 9, 10 Hz).

The supersonic molecular beam of salicin was further collimated through a skimmer of 2 mm diameter and intersected with a frequency doubled output (0.2 – 0.3 mJ) of a tunable dye laser (ND6000, Continuum) pumped by second harmonic output of a Nd:YAG laser (10

nanosecond, 10 Hz, Surelite II-10, Continuum). The electronic spectrum of salicin was measured using one-color resonant 2-photon ionization (1C-R2PI) spectroscopy. The desorption laser and the ionization laser were delayed by 160  $\mu$ s and 400  $\mu$ s, respectively, from the firing of the pulsed valve driver (IOTA ONE, Parker).

UV-UV hole-burning spectroscopy was used to discriminate the presence of different conformers of salicin in the experiment. In this technique, two counter-propagating UV lasers were used as pump and probe lasers. The probe laser (10 nanosecond, 10 Hz, Surelite II-10, Continuum) was fixed at the wavelength of one of the R2PI peaks while the pump laser (10 nanosecond, 10 Hz, Surelite II-10, Continuum) was scanned through the R2PI spectral region of the molecule. The pump laser preceded the probe laser by about 100 ns. The hole-burning spectrum provided depletion in the ion signal of the R2PI peaks which belonged to the same conformer.

IR spectra of different conformers of salicin were recorded using resonant ion-dip infrared spectroscopy (RIDIRS). In RIDIRS, the UV laser was fixed at a particular R2PI peak while the IR laser, which preceded the UV laser by  $\sim$  100 ns, was scanned in the O-H stretching frequency region. The IR spectrum of a particular conformer was obtained as a depletion in the R2PI signal whenever the IR laser frequency was in resonance with a vibrational transition. The IR laser (Laser Vision, pulse energy  $\sim$  4-5 mJ, resolution  $\sim$  2.5  $\text{cm}^{-1}$ ) was pumped by fundamental output of an unseeded Nd:YAG laser (Continuum, Surelite II-10, 10 nanosecond, 10 Hz). Synchronization between the lasers and the pulsed valve was performed using a digital delay generator (BNC 575). Home-built LabVIEW based programs were used for data acquisition and laser scanning.

## 2.2. Computational Methods

Many conformations of salicin are possible due to its multiple flexible co-ordinates. Firstly, a conformational search program CONFLEX<sup>56-58</sup> based on a force field calculation (MMFF94s) was used to generate probable initial structures in the conformational space of salicin. The structures within a relative energy cut-off of  $\sim$ 9 kcal/mol with respect to the lowest energy structure were considered for geometry optimization using density functional theory (DFT) calculations. Geometry optimization and vibrational frequency calculations of various structures

of salicin in the ground state ( $S_0$ ) were performed using the M06-2X density functional,<sup>59</sup> which has recently been demonstrated to be the best performing of the Minnesota density functionals for non-covalent interactions,<sup>60</sup> using the Gaussian suite of programs.<sup>61</sup> The optimized geometries and vibrational frequencies of a few selected conformers in the excited state ( $S_1$ ) were obtained using the time-dependent (TD) M06-2X method, which has been demonstrated to be one of the most accurate hybrid functionals for TD calculation relative to both experimental and theoretical reference data.<sup>62, 63</sup> The 6-311++G(d,p) basis set was used unless otherwise stated, along with a pruned integration grid of 99 radial shells and 590 angular points per shell (the so-called “UltraFine” grid). Vibrationally-resolved electronic spectra were calculated in a Franck-Condon simulation using Gaussian with  $S_0$  and  $S_1$  structures and vibrational frequencies calculated at the (TD-)M06-2X/6-311++G(d) levels of theory.<sup>64</sup> As the Hessian for the  $S_1$  state is calculated numerically the polarization functions were removed from hydrogen atoms to keep calculation times reasonable, with initial testing showing this led to a negligible effect on ground state geometries. The Franck-Condon spectra were also inspected using the FC-LabII program<sup>65</sup> for the purposes of band assignment.

Relative Gibbs free energies ( $\Delta G$ ) of various conformers of salicin were calculated at 10 K using a numerical program available on the NIST website to compute ideal gas thermodynamic functions at different temperatures.<sup>49, 66</sup> This program uses the thermodynamic parameters obtained from Gaussian calculations at 298 K.  $\Delta G$  has been computed using the equation  $\Delta G = E_{\text{elec}} + E_{\text{ZPE}} + E_{\Delta H} - T\Delta S$ ; where  $E_{\text{elec}}$  is total electronic energy of molecules,  $E_{\text{ZPE}}$  is zero-point energy,  $\Delta H$  is enthalpy change,  $T$  is temperature and  $\Delta S$  is entropy change. Natural bond orbital (NBO) calculations<sup>67</sup> were carried out using the NBO6.0 program with the M06-2X/6-31++G(d,p) density functional theory.<sup>68</sup>

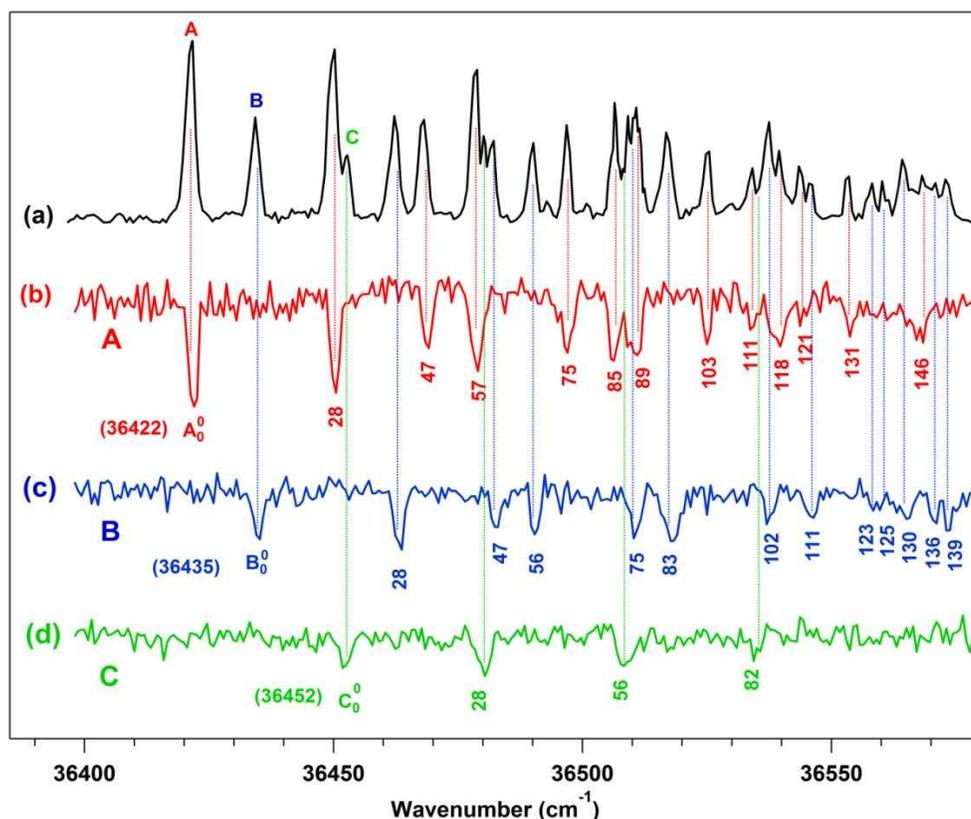
SAPT calculations<sup>69</sup> have been carried out on several low energy conformations of salicin using the PSI4 package.<sup>70</sup> The underlying density fitted SAPT0 method used the cc-pVDZ<sup>71</sup> basis set for hydrogen atoms and the diffuse augmented aug-cc-pVDZ<sup>72</sup> basis with the diffuse d-type functions removed for all heavier atoms. For a functional group SAPT (F-SAPT) analysis the salicin molecule was partitioned into two localized chemical functional groups: the benzylic OH and the phenyl ring. The remaining atoms in the molecule formed a linking unit, such that the F-SAPT interaction was carried out within the embedding field of said linking unit. Further

details on this F/I-SAPT procedure are detailed elsewhere.<sup>73, 74</sup> The effect of dispersion on the ground state equilibrium geometries of salicin conformers was carried out by comparing B3LYP<sup>75, 76</sup> and B3LYP-D (D3 dispersion correction with Becke-Johnson damping)<sup>77</sup> geometries, both using the 6-311++G(d,p) basis along with an UltraFine integration grid in the Gaussian package. Root mean square deviations (RMSDs) of the geometries were calculated using the Quaternion algorithm for rotation.<sup>78, 79</sup>

## Results and discussion

### Conformations of salicin: Electronic spectroscopy

The electronic spectrum of salicin measured by 1C-R2PI spectroscopy is shown in Figure 2(a). The spectrum shows several sharp bands in the range of 36400 – 36600  $\text{cm}^{-1}$ . To determine whether all the electronic bands shown in Figure 2(a) correspond to single or multiple



**Figure 2.** (a) Electronic spectrum of salicin measured using 1C-R2PI spectroscopy. (b), (c) and (d) are UV-UV hole burning spectra of salicin measured by probing the bands labeled as A (36422), B (36435) and C (36452  $\text{cm}^{-1}$ ) respectively, in the R2PI spectrum. The electronic band origins of A, B and C conformers are designated as  $A_0^0$  (36422  $\text{cm}^{-1}$ ),  $B_0^0$  (36435  $\text{cm}^{-1}$ ) and  $C_0^0$  (36452  $\text{cm}^{-1}$ ).

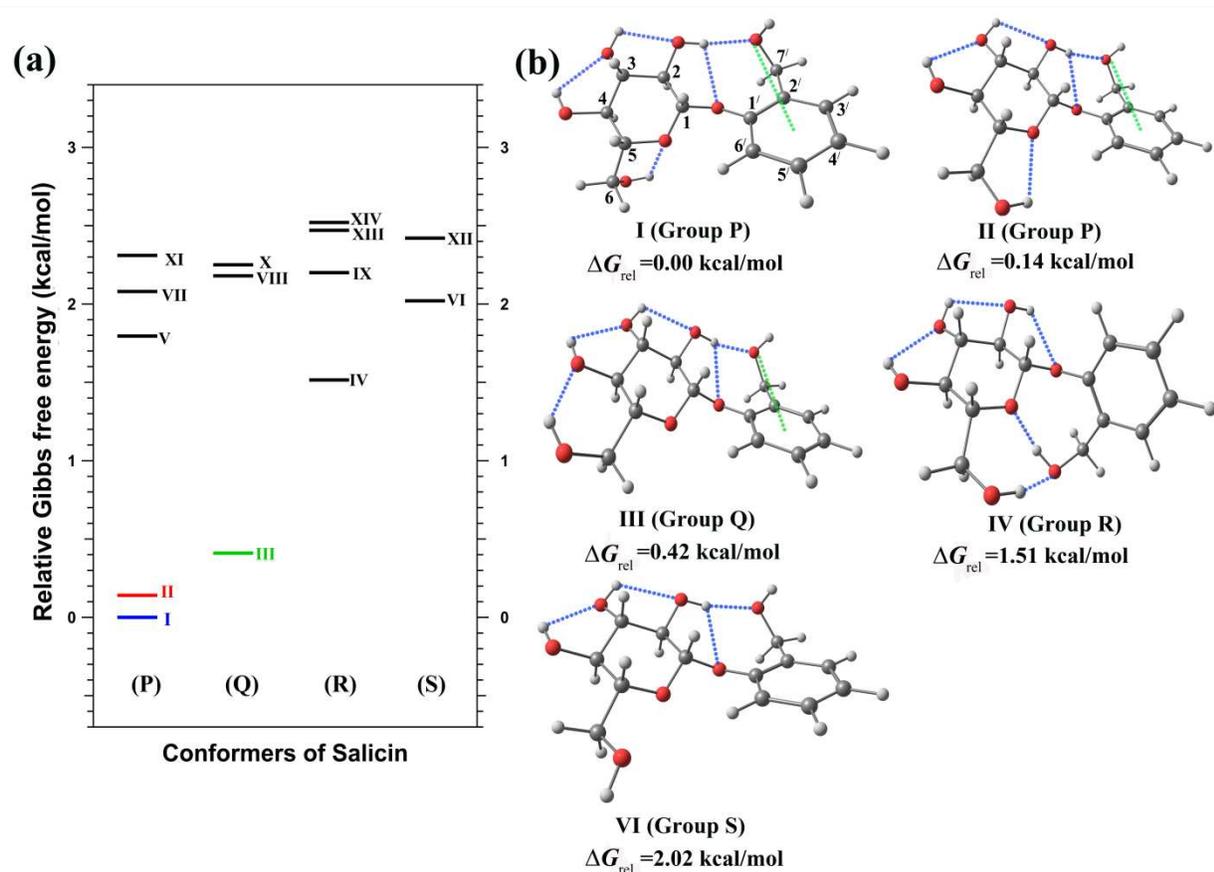
conformers of salicin, UV-UV hole-burning spectroscopy was performed. Figures 2(b), (c) and (d) show UV-UV hole-burning spectra measured by probing the bands labeled as A, B, and C, respectively, in the spectrum displayed in Figure 2(a). In general, a hole-burning spectrum probing a particular band in the electronic spectrum provides depletion of ion signals for all the electronic bands which belong to a specific conformer. The spectra shown in Figures 2(b), 2(c) and 2(d) clearly show the presence of three distinct conformers of salicin in the experiment. The hole-burning spectra shown in Figures 2(b), 2(c) and 2(d) are basically conformation-specific electronic spectra of conformers A, B, and C, respectively, of salicin. The  $S_1 \leftarrow S_0$  origin bands ( $0_0^0$ ) of conformers A, B, and C labeled as  $A_0^0$ ,  $B_0^0$  and  $C_0^0$  appear at 36422, 36435, and 36452  $\text{cm}^{-1}$ , respectively. The intensity pattern of the electronic bands of the three conformers reveals that the most populated conformer is A, while the C conformer is least populated. All three conformers show a progression of a low-frequency vibrational mode of 28  $\text{cm}^{-1}$ . The electronic spectra of conformers A and B show some other low frequency modes (e.g. 47  $\text{cm}^{-1}$ ) and their combination bands apart from the 28  $\text{cm}^{-1}$  mode. The detailed assignment of the electronic bands of the three conformers made through their  $S_1$  state vibrational frequency calculations and simulated Franck-Condon spectra is provided below and in the Supporting Information (SI).

As salicin is a close analog of phenyl  $\beta$ -D-glucopyranoside having additional substitution of a hydroxymethyl group at 2-position of the phenyl group, it is worth comparing the electronic spectrum of salicin with that of the phenyl  $\beta$ -D-glucopyranoside reported in the literature.<sup>47</sup> It is interesting to note that the electronic spectrum of phenyl  $\beta$ -D-glucopyranoside also indicates the presence of three conformers in the experiment. The low frequency vibronic bands observed in the electronic spectra of the three conformers of phenyl  $\beta$ -D-glucopyranoside are also similar to those of salicin.

### **Conformational landscape of salicin: DFT calculations**

Although only three conformers are observed in the experiment, salicin can have many possible conformations due to its flexible shape. Thus proper assignment of the structures of the observed conformers is not straightforward. Firstly, the CONFLEX program was used to generate probable initial structures in the conformational space of salicin. A total of 60 conformers were generated from the force field calculation and 35 structures were selected from there on the basis of an energy cut-off of  $\sim 9$  kcal/mol relative to the most stable conformer for

quantum chemical calculations. Structures of all the 35 conformers were optimized at the M06-2X/6-311++G(d,p) level of theory. Table S1 of the supporting information shows the comparison of relative energies of all the 35 conformers of salicin obtained from the force field calculation and those obtained from quantum chemical calculation. Only the 14 lowest energy conformers have been considered for further investigations as these conformers have energies within 3 kcal/mol of the lowest energy conformer, and the higher energy conformers are less likely to be observed in the experiment. The optimized structures of these 14 conformers were then arranged in order of increasing energy from the global minimum as well as into different groups according to their structural similarities.



**Figure 3.** (a) Relative Gibbs free energies of the conformers of salicin calculated [M06-2X/6-311++G(d,p)] at 10K. Conformers are divided into four groups P, Q, R, and S on the basis of their structural similarities. (b) Representative structures from groups P, Q, R and S.

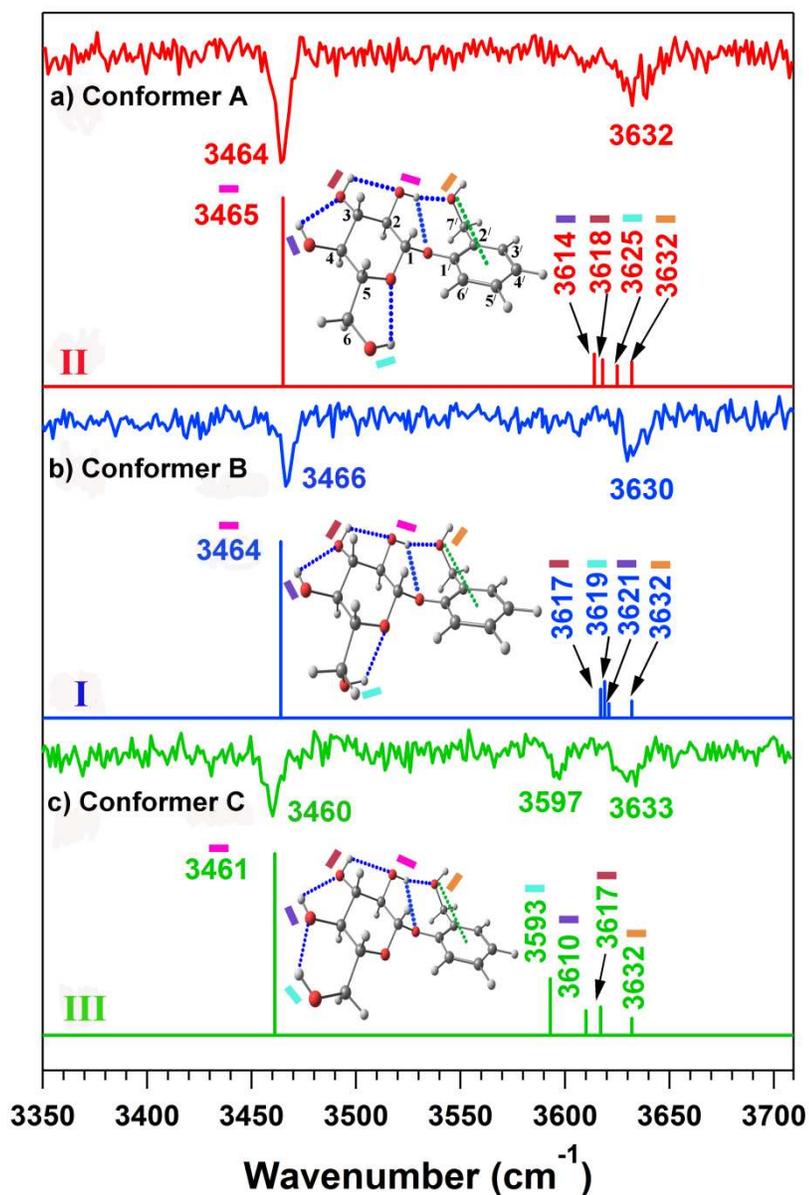
Relative Gibbs free energies ( $\Delta G_{rel}$ ) of the 14 low energy conformers calculated at 10K and their classification into four groups (P, Q, R, S) in terms of the structural similarities are

shown in Figure 3(a). For brevity, only a representative structure from each of the four groups has been provided in Figure 3(b), while all 14 structures are provided in Figure S1 in the Supporting Information (SI). The second lowest energy conformer (II) which belongs to the P group is also presented in Figure 3(b) as it is energetically very close to the global minimum. It is observed that major structural changes of these conformations occur mostly due to the change in the orientation of the hydroxy methyl (-CH<sub>2</sub>OH) groups present in the pyranose and phenyl groups of salicin. The structural details of the conformers in the four groups have been visualized in Figure S1 of the SI. The general structural motif of all the conformers is a chain of hydrogen bonds among the -OH groups of the pyranose ring and benzyl alcohol moiety. The structural details of the conformers are also discussed in the SI.

The most interesting feature of Figure 3(a) is that the three lowest energy conformers of salicin are conformers I, II, and III, which are within 0.5 kcal/mol relative Gibbs free energy ( $\Delta G_{\text{rel}}$ ). The next highest energy conformer (conformer IV) is comparatively high in energy with a  $\Delta G_{\text{rel}}$  value of roughly 1.5 kcal/mol. As three conformers of salicin are observed in the experiment, conformers I, II, and III are probable candidates for the observed conformers. Similar trends in the relative energies of various conformers of salicin were obtained with other density functionals and basis sets, and these results have been listed in Table S2 in the SI. Comparison between relative electronic energies ( $E_{\text{rel}}$ ) and relative Gibbs free energies ( $\Delta G_{\text{rel}}$ ) of all the conformers have been provided in Table S3 in the SI. It has been found that the trend in  $E_{\text{rel}}$  and  $\Delta G_{\text{rel}}$  of the conformers are quite similar. Thus, despite salicin being a flexible molecule and having an abundance of conformers, the determination of the structures of the experimentally observed conformers becomes relatively straightforward when paired with the output of quantum chemical calculations.

### **Structures of observed conformers of salicin: IR spectroscopy**

The primary aim of the present work is to determine the structures of the conformers of salicin observed in the experiment and understand the interplay between the non-covalent interactions that govern their energetics. Figures 4(a), 4(b), and 4(c) show experimental IR spectra of the three conformers (marked as A, B, and C in the R2PI spectrum of Figure 2) of salicin in the O-H stretching region measured by probing their respective electronic origin band using RIDIR



**Figure 4.** (a), (b) and (c) show the experimental IR spectra of species A, B and C, respectively, in the O-H stretching region measured by probing their respective electronic band origins (for electronic origin bands see Figure 2). The scaled theoretical O-H stretching frequencies of conformers II, I and III shown as stick spectra are assigned to conformers A, B and C respectively. The theoretical O-H stretching frequencies are calculated at the M06-2X/6-311++G(d,p) level of theory. The structures of conformers I, II, and III are also shown in the inset of the figure. The blue dotted line represents the O-H...O hydrogen bonding and the green dotted line represents the  $n \rightarrow \pi^*$  interaction.

spectroscopy. As there are a total of five OH groups in salicin, five bands should be observed in the O-H stretching frequency region of the IR spectra if all the bands are well resolved. The IR spectra of all three conformers show a strong band around the  $3460 \text{ cm}^{-1}$  region. On the other

hand, the remaining four IR bands appearing in the 3590-3640  $\text{cm}^{-1}$  region are quite weak, broad and unresolved. Although the position of the strong IR band is very close across all three conformers, their respective positions (3464, 3466, and 3460  $\text{cm}^{-1}$ ) have been confirmed through repeated IR scans. Conformers A and B show the weak broad features centering around 3630  $\text{cm}^{-1}$  while conformer C shows a distinct band at 3597  $\text{cm}^{-1}$  in addition to the broad feature at 3630  $\text{cm}^{-1}$ . Thus, it is apparent that the structures of conformers A and B are quite similar, while the structure of conformer C is a little different compared to the other two.

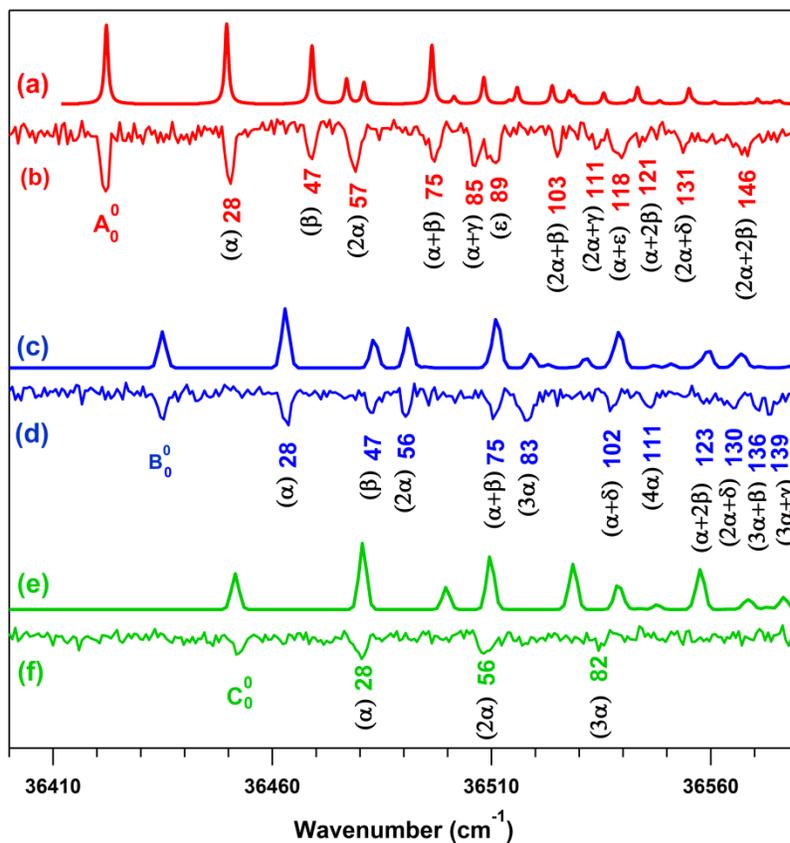
Experimental IR spectra of these three conformers are then compared with the theoretical IR spectra of the conformers classified into P, Q, R, and S groups computed at the M06-2X/6-311++G(d,p) level of theory. The calculated O-H harmonic stretching frequencies of all fourteen conformers are scaled using a factor of 0.9348, which is the ratio of the experimental to calculated OH stretching frequency of ethanol<sup>80</sup> at the same level of theory. Similar scaling factors have been obtained with respect to experimental OH stretching frequency of benzyl alcohol<sup>81</sup> and methanol<sup>82</sup> (see Table S4 in the SI). The interesting point here is that the pattern of the experimental OH stretching frequencies serves as a guide to choosing probable observed conformers from the pool of structures obtained from the series of calculations. Theoretical IR spectra of the three lowest energy conformers (I, II and III) of salicin along with their computed structures are shown in Figure 4 while the comparison of the experimental IR spectra of the three conformers of salicin with the theoretical IR spectra of all 14 conformers is provided in Figure S2 of the SI. Each of the five O-H groups in the structures and the corresponding IR band in the theoretical IR spectra (Figure 4) are marked with a specific color bar. It is clear that the C(2)-OH group of the sugar moiety is very strongly hydrogen bonded to the oxygen atom of the OH group of the benzyl alcohol moiety, while the hydrogen bonds involving other OH groups of the sugar moiety are very weak and similar in nature. It can be concluded from the comparison shown in Figure 4 that the IR spectra of conformers A and B are due to structures I and II. However, it is not straightforward to tell whether conformer A has structure I and conformer B has structure II, or vice versa. On the other hand, conformer C can easily be assigned to structure III from the comparison of the experimental IR spectrum provided in Figure 4(c) with the theoretical IR spectrum of III, as the IR spectrum of conformer C is different from those of conformers A and B in the 3590-3650  $\text{cm}^{-1}$  region. In structure III, the  $-\text{CH}_2\text{OH}$  group of the sugar moiety is hydrogen bonded to the neighboring OH group, while the  $-\text{CH}_2\text{OH}$  group of the sugar moiety in

both structures I and II is hydrogen bonded to the oxygen atom of the sugar ring. The IR band of conformer C at  $3597\text{ cm}^{-1}$  [Figure 4(c)] is assigned to the OH vibration of the  $-\text{CH}_2\text{OH}$  group of the sugar moiety. Theoretical IR spectra of the remaining higher energy conformers (structures IV-XIV) presented in Figure S2 can be disregarded based on either the higher relative energy or the significantly different IR spectra of these conformers compared to the three lowest energy conformers (I, II, III).

It is intriguing to compare the IR spectra of the three conformers of salicin with those of phenyl  $\beta$ -D-glucopyranoside reported by Simons and co-workers.<sup>47</sup> It can be seen that the IR spectra of the two major conformers of phenyl  $\beta$ -D-glucopyranoside are similar to those of conformers A and B of salicin in the broad and weak spectral region ( $3620\text{-}3650\text{ cm}^{-1}$ ). The IR spectrum of the minor conformer of phenyl  $\beta$ -D-glucopyranoside matches well with that of the conformer C of salicin in the  $3590\text{-}3650\text{ cm}^{-1}$  region. However, the strong IR band of salicin around  $3460\text{ cm}^{-1}$  is absent in the IR spectra of all conformers of phenyl  $\beta$ -D-glucopyranoside. The comparison of the IR spectra of salicin and phenyl  $\beta$ -D-glucopyranoside thus indicates that the strong band ( $3460\text{-}3466\text{ cm}^{-1}$ ) observed in the IR spectrum of salicin could be due to a strong hydrogen bonding interaction between the benzylic  $\text{CH}_2\text{OH}$  group and  $\text{O}2\text{H}2$  group. On the other hand, the weak unresolved broad peaks are due to the OH groups in the sugar moiety present as a chain of intramolecular  $\text{O-H}\dots\text{O}$  hydrogen bonds. The hydrogen-bonded structural motifs of the three observed conformers of phenyl  $\beta$ -D-glucopyranoside are quite similar to those of the three observed conformers of salicin. Thus the comparison between salicin and phenyl  $\beta$ -D-glucopyranoside data provides extra confidence to the assignment of the structures of the observed salicin conformers. It should also be noted that the substitution of the  $-\text{CH}_2\text{OH}$  group in the phenyl moiety of the sugar derivative does not change the basic structural motif of the sugar unit.

### **Franck-Condon simulated electronic spectra**

As mentioned earlier, it is difficult to predict which of the two lowest energy conformers (I and II) of salicin belongs to species A and B in the electronic spectrum due to the similarity in their IR spectra. However, Franck-Condon (FC) simulated electronic spectra of different conformers of salicin and their vertical excitation energies can further aid in assigning the structures of the observed conformers.<sup>83, 84</sup> Figures 5(b), (d), and (f) show experimentally observed electronic



**Figure 5.** (b), (d) and (f) shows experimental electronic spectra of species A, B and C respectively, in comparison with the Franck-Condon simulated electronic spectra of conformers II, I and III presented in (a), (c) and (e), respectively. Assignments of the bands are based on the simulated Franck-Condon spectra.

spectra of species A, B, and C, respectively, while FC simulated electronic spectra of conformers II, I, and III are shown in Figures 5(a), (c), and (e), respectively. There is an excellent agreement (in terms of the frequency of vibration as well as the intensity of the bands) between the experimentally observed and FC simulated electronic spectra of the three conformers of salicin. It can be seen that the origin band of conformer B (B<sub>0</sub><sup>0</sup>) is weaker in intensity when compared to the 28 cm<sup>-1</sup> vibration of conformer B, while for conformer A the intensities of the origin band (A<sub>0</sub><sup>0</sup>) and 28 cm<sup>-1</sup> vibration are comparable in magnitude. It is noteworthy that this specific intensity pattern of experimental electronic spectra of species A and B corroborates that of the FC simulated electronic spectra of conformers II and I, respectively. Thus species A and B can be assigned as conformers II and I, respectively. Vertical excitation energies of the three lowest energy conformers (I, II, and III) calculated at the TD-M06-2X/6-311++G(d) level of theory lend

additional support to the assignment of the structures of the conformers observed in the experiment. Vertical excitation energies (scaled with respect to the origin band of species A) of conformers I, II, and III of salicin are provided in Table S6 of the SI. Conformer II has lowest vertical excitation energy. The order of relative vertical excitation energies of conformers II, I, and III matches well with the electronic origin band positions of species A, B, and C, respectively. This suggests that species A, B, and C should be assigned as conformers II, I, and III, respectively. FC simulated electronic spectra of a few higher energy conformers of salicin were also calculated. However, large geometric changes between  $S_0$  and  $S_1$  states mean the calculated spectra have almost zero intensity.

Low frequency vibronic bands present in the electronic spectra of the three conformers of salicin are assigned with the help of the simulated Franck-Condon spectra (see Figure 5). A list of the experimental and calculated ( $S_1$ ) low-frequency vibrational modes (based on FC simulation) as well as a tentative assignment of all the low frequency bands in the electronic spectra of the three lowest energy conformers of salicin is also provided in Table S7 of the SI. It is found that there is a long progression of the  $28\text{ cm}^{-1}$  vibration ( $\alpha$ ) in the electronic spectra of all three conformers. This  $28\text{ cm}^{-1}$  vibration is assigned as an inter-ring twisting vibration along the  $\angle\text{O7-C1-O8-C1}'$  dihedral angle. Another prominent low frequency vibrational mode of  $47\text{ cm}^{-1}$  is observed for both conformers A and B while the same is not observed for conformer C. The electronic spectra of A and B can be assigned mostly by the overtones of the  $\alpha$  mode and its combination bands with other modes. Thus, the interpretation of the electronic spectra suggests that major geometrical changes in salicin after electronic excitation may occur along this  $\alpha$  mode.

### **Conformational preference in salicin: $n\rightarrow\pi^*$ interaction**

The structures of the conformers I, II and III calculated at the M06-2X/6-311++G(d,p) level of theory are provided in Figure 3b and Figure 4, with their important geometrical parameters and full geometries (in Cartesian coordinates) presented in the SI. It is evident that all three observed conformers differ mainly in the orientation of the hydroxy group (O6-H6) of the -CH<sub>2</sub>OH group in the sugar moiety, which is represented by the dihedral angle O6-C6-C5-O5. In both conformers I and II, O6-H6 has a gauche orientation with  $\angle\text{O6-C6-C5-O5}$  of  $+57^\circ/-57^\circ$ .

The O6-H6 group of conformer I and II is involved in a weak hydrogen bonding interaction with the O5 atom of the pyranose ring (O6-H6...O5). The O6-H6...O5 hydrogen bond distance and angle are around 2.3 Å and 105°, respectively, in both conformers I and II. Conformer III differs from I and II as the O6-H6 group in the former has a trans orientation ( $\angle\text{O6-C6-C5-O5} = 164.5^\circ$ ). Unlike conformers I and II, the O6-H6 group of conformer III is involved in a hydrogen bonding interaction with the O4-H4 group (O6-H6...O4). Purely based on a hydrogen bond distance and angle point of view<sup>85</sup> (see Table S8 in the SI), the O6-H6...O4 hydrogen bond in conformer III is stronger than the O6-H6...O5 hydrogen bond present in I and II. We have analyzed here H...O distance to indicate the strength of the hydrogen bond.

In all three conformers, the equatorial hydroxy groups of the sugar moiety adopt the gauche orientation. Equatorial OH groups are linked with each other through a chain of weak O-H...O hydrogen bonds. This is evident through their hydrogen bond distances and angles, which are around 2.5 Å and 106°, respectively, in all observed conformers. However, all three conformers (I, II and III) have a strong intramolecular hydrogen bond between the hydroxy group (O2-H2) of the sugar moiety and the oxygen atom (O7') of the benzyl alcohol unit (O2-H2...O7'). Hydrogen bonding interactions are usually stronger when the hydrogen bond angle ( $\angle\text{X-H...Y}$ ) is close to 180° and this is accompanied by a shorter hydrogen bond distance.<sup>85</sup> The O2-H2...O7' hydrogen bond distance (~1.9 Å) and angle (~164°) lie close to the ideal geometrical conditions required for strong hydrogen bonding interaction<sup>85</sup> in all three conformers. This indicates that the O2-H2...O7' hydrogen bond is much stronger than the hydrogen bonds between the equatorial OH groups.

Although only the three lowest energy conformers (I, II, and III) are observed in the experiment, analysis of the hydrogen bonding parameters in a few higher energy conformers reveals interesting information that is the focus of the present section. From Figure 3(a) it can be seen that conformer V is the next highest energy conformer after I and II in group P, while conformer VIII is the next highest energy conformer after III in group Q. It is evident from the comparison of the geometrical parameters of the conformers presented in Table S8 that the O2-H2...O7' hydrogen bond in both conformers V and VIII ( $\angle\text{O2-H2...O7}' \sim 172^\circ$ ) is stronger than that in conformers I, II, and III ( $\angle\text{O2-H2...O7}' \sim 162^\circ$ ). This suggests that the other O-H...O hydrogen bond interactions in conformers V and VIII are similar in strength to those of I/II and

III. In spite of this, it is interesting to note that conformers V and VIII are relatively high in energy compared to conformers I, II, and III. Basically, conformer V belongs to the same group of conformers as both I and II (group P in Figure 3), while conformer VIII is in group Q with conformer III.

It is apparent from the structures of salicin provided in Figure S1 and discussion of the structural details of the conformers in the SI that the structures of conformers V and VIII differ from those of conformers I/II and III, respectively, due solely to the orientation of the O<sub>7</sub>'-H<sub>7</sub>' group of the benzyl alcohol moiety of salicin. In all three observed conformers of salicin, i.e. conformers I, II, and III, the O<sub>7</sub>'-H<sub>7</sub>' group of the benzyl alcohol unit has gauche orientation ( $\angle\text{H7}'\text{-O7}'\text{-C7}'\text{-C2}' \sim 56.8^\circ$ ) and the distance of the O<sub>7</sub>'-H<sub>7</sub>' group from the center of the aromatic ring ( $r_{\text{O7}'\text{-H7}'\text{...Ar}}$ ) is  $\sim 3.6 \text{ \AA}$ . In fact, the gauche orientation of the O<sub>7</sub>'-H<sub>7</sub>' group of conformer I, II and III resembles the orientation of the OH group in the gauche conformer of isolated benzyl alcohol.<sup>81</sup> On the other hand, conformers V and VIII have a trans orientation of the O<sub>7</sub>'-H<sub>7</sub>' group ( $\angle\text{H7}'\text{-O7}'\text{-C7}'\text{-C2}' \sim 167^\circ$ ) which resembles the orientation of the OH group in the trans conformer of isolated benzyl alcohol. The gauche conformer of benzyl alcohol is reported to be more stable than the trans conformer based on both experimental and theoretical data.<sup>81</sup> Interestingly, conformers I, II, and III of salicin (gauche orientation of the O<sub>7</sub>'-H<sub>7</sub>' group) are also more stable than conformers V and VIII of salicin, which have trans orientation of the O<sub>7</sub>'-H<sub>7</sub>' group. Based on this, it appears that the gauche orientation of the benzylic OH group plays an important role in governing the conformational preferences of salicin.

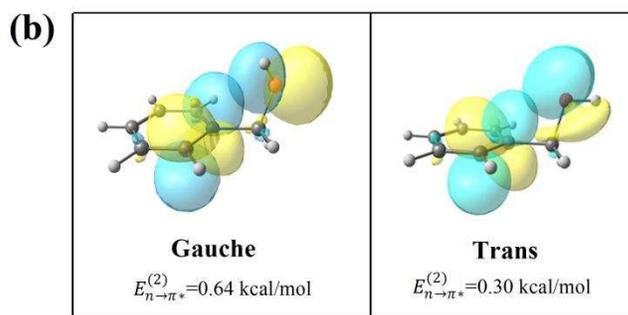
The stability of the gauche conformer of benzyl alcohol over the trans conformer has previously been rationalized by the presence of an O-H... $\pi$  interaction in the former.<sup>81</sup> However, an NBO analysis on the gauche conformer of benzyl alcohol, as well as conformers I, II, and III of salicin having gauche orientation of the benzylic OH group, does not reveal any signature of an O-H... $\pi$  interaction (see Figure S3 of the SI). Generally, the presence of an O-H... $\pi$  interaction is manifested through delocalization of electrons in the  $\pi$ -orbitals to the antibonding  $\sigma_{\text{O-H}}$  orbital. On the contrary, NBO analysis reveals that the lone pair orbitals (both sp and p-types) on the oxygen atom of the OH group of the gauche conformer of benzyl alcohol, as well as for salicin conformers I, II, and III, have significant delocalization with the  $\pi^*$  orbitals of the phenyl group and this non-covalent interaction is termed  $n \rightarrow \pi^*$ .<sup>4</sup> It should be noted that the

oxygen atom has two lone pair orbitals and generally, one of the lone pair orbitals has maximum ‘p’ character and thus is denoted ‘p’ type orbital while the other lone pair orbital having mixed ‘s’ and ‘p’ character is denoted ‘sp’ type. NBOs for the  $n \rightarrow \pi^*$  interaction as well as the O2-H2...O7’ hydrogen bonding interaction in conformers I, II, III, V and VIII of salicin are shown in Figure 6(a). NBOs for the  $n \rightarrow \pi^*$  interaction in gauche and trans conformers of benzyl alcohol are also shown in Figure 6(b). NBO second-order perturbative estimates of donor-acceptor interactions ( $E_{i \rightarrow j}^{(2)}$ ) for all the O-H...O hydrogen bonds and  $n \rightarrow \pi^*$  interactions in conformers I, II, III, V and VIII of salicin are listed in Table 1.

**Table 1:** NBO second-order perturbative estimates of donor-acceptor interactions ( $E_{i \rightarrow j}^{(2)}$ ) of all hydrogen bonding and  $n \rightarrow \pi^*$  interactions observed in conformers I, II, III, V and VIII of salicin. NBO calculations were performed at the M06-2X/6-311++G(d,p) level of theory.

	$E_{i \rightarrow j}^{(2)}$ (kcal/mol)				
	I	II	III	V	VIII
$n_p(O7') \rightarrow \sigma^*(O2-H2)$	5.65	5.35	5.90	7.07	7.61
$n_{sp}(O7') \rightarrow \sigma^*(O2-H2)$	3.28	3.24	3.27	2.39	2.40
$n_p(O2) \rightarrow \sigma^*(O3-H3)$	0.15	0.22	0.22	0.33	0.32
$n_{sp}(O2) \rightarrow \sigma^*(O3-H3)$	0.21	0.16	0.15	0.18	0.18
$n_p(O3) \rightarrow \sigma^*(O4-H4)$	0.15	0.20	0.20	0.23	0.22
$n_{sp}(O3) \rightarrow \sigma^*(O4-H4)$	0.18	0.20	0.24	0.20	0.24
$n_p(O5) \rightarrow \sigma^*(O6-H6)$	0.18	0.16	-	-	0.10
$n_{sp}(O5) \rightarrow \sigma^*(O6-H6)$	0.17	0.15	-	0.15	0.14
$n_p(O4) \rightarrow \sigma^*(O6-H6)$	-	-	0.69	-	0.68
$n_{sp}(O4) \rightarrow \sigma^*(O6-H6)$	-	-	2.60	-	2.61
$n_p(O7') \rightarrow \pi^*(C2'/-C3')$	0.80	0.83	0.78	-	-
$n_{sp}(O7') \rightarrow \pi^*(C2'/-C3')$	0.33	0.32	0.35	0.31	0.31

	Hydrogen bond		$n \rightarrow \pi^*$	
	$O_{7'}(n_p) \rightarrow \sigma^*(O_2 - H_2)$	$O_{7'}(n_{sp}) \rightarrow \sigma^*(O_2 - H_2)$	$O_{7'}(n_p) \rightarrow \pi^*(C_2=C_3)$	$O_{7'}(n_{sp}) \rightarrow \pi^*(C_2=C_3)$
<b>I</b>	 $E_{n \rightarrow \sigma^*}^{(2)} = 5.65$ kcal/mol	 $E_{n \rightarrow \sigma^*}^{(2)} = 3.28$ kcal/mol	 $E_{n \rightarrow \pi^*}^{(2)} = 0.80$ kcal/mol	 $E_{n \rightarrow \pi^*}^{(2)} = 0.33$ kcal/mol
<b>II</b>	 $E_{n \rightarrow \sigma^*}^{(2)} = 5.35$ kcal/mol	 $E_{n \rightarrow \sigma^*}^{(2)} = 3.24$ kcal/mol	 $E_{n \rightarrow \pi^*}^{(2)} = 0.83$ kcal/mol	 $E_{n \rightarrow \pi^*}^{(2)} = 0.32$ kcal/mol
<b>III</b>	 $E_{n \rightarrow \sigma^*}^{(2)} = 5.90$ kcal/mol	 $E_{n \rightarrow \sigma^*}^{(2)} = 3.27$ kcal/mol	 $E_{n \rightarrow \pi^*}^{(2)} = 0.78$ kcal/mol	 $E_{n \rightarrow \pi^*}^{(2)} = 0.35$ kcal/mol
<b>V</b>	 $E_{n \rightarrow \sigma^*}^{(2)} = 7.07$ kcal/mol	 $E_{n \rightarrow \sigma^*}^{(2)} = 2.39$ kcal/mol	—	 $E_{n \rightarrow \pi^*}^{(2)} = 0.31$ kcal/mol
<b>VIII</b>	 $E_{n \rightarrow \sigma^*}^{(2)} = 7.61$ kcal/mol	 $E_{n \rightarrow \sigma^*}^{(2)} = 2.40$ kcal/mol	—	 $E_{n \rightarrow \pi^*}^{(2)} = 0.31$ kcal/mol



**Figure 6** (a) NBOs of the  $O_2-H_2 \dots O_{7'}$  hydrogen bonding interaction and  $n \rightarrow \pi^*$  interaction in various conformers (I, II, III, V and VIII) of salicin calculated at the M06-2X/6-311++G(d,p) level of theory. The NBOs of the  $n \rightarrow \pi^*$  interaction show delocalization of electrons from the lone pair orbitals ( $n_{sp}$  and  $n_p$ ) of  $O_{7'}$  atom to the  $\pi^*$  orbital of  $C_2=C_3$  bond of the aromatic ring. The NBOs of the hydrogen bonding interaction show delocalization of electrons from lone pair orbitals ( $n_{sp}$  and  $n_p$ ) of  $O_{7'}$  atom to the  $\sigma^*$  orbital of the  $O_2-H_2$  bond. (b) NBOs of the  $n \rightarrow \pi^*$  interaction between the lone pair orbital of ( $n_{sp}$  and  $n_p$ ) oxygen atom and the aromatic ring in the gauche and trans conformers of benzyl alcohol.

The sum of the  $E_{i \rightarrow j}^{(2)}$  values for the  $n \rightarrow \pi^*$  interactions in the gauche conformer of benzyl alcohol and conformers I, II, and III of salicin is about 1.2 kcal/mol. The trans conformer of benzyl alcohol, as well as conformers V and VIII of salicin, also have an  $n \rightarrow \pi^*$  interaction, albeit significantly weaker at around 0.4 kcal/mol (see below). Interestingly, NBO derived estimates for the strong hydrogen bond O2-H2...O7' in different conformers of salicin listed in Table 1 also highlight that the O2-H2...O7' hydrogen bond in conformers V and VII of salicin is slightly stronger than that in conformers I, II, and III. The data on the relative strength of different hydrogen bonds in various conformers of salicin obtained from NBO calculations are also supported by quantum theory of atoms in molecules (QTAIM) analysis which has been provided in Table S9 in the SI. Thus it could be concluded that conformational preferences of conformers I, II, and III over conformers V and VIII of salicin might be due to a subtle interplay between the hydrogen bond (O2-H2...O7') and the  $n \rightarrow \pi^*$  interaction present in the former conformers. Readers are reminded that the sum of the total  $E_{i \rightarrow j}^{(2)}$  values obtained from NBO calculations does not represent the total interaction energy of the system.

A further interesting point from Table 1 is that both sp- and p-type lone pair electrons ( $n_{sp}$  and  $n_p$ ) on the O7' atom of the benzylic O-H group in conformers I, I, and III of salicin simultaneously participate in hydrogen bonding as well as the  $n \rightarrow \pi^*$  interaction. The  $n \rightarrow \pi^*$  interaction involving the  $n_p$ -type electrons is stronger than that involving the  $n_{sp}$ -type electrons on the O7' atom in conformers I, II, and III of salicin. On the other hand, the  $n_p$ -type electrons on the O7' atom in conformers V and VIII do not take part in the  $n \rightarrow \pi^*$  interaction. As the  $n_p$ -type electrons on O7' atom are involved in the stronger  $n \rightarrow \pi^*$  interaction in conformers I, II and III, the extent of delocalization of this lone pair to the  $\sigma^*$  orbital of the O2-H2 bond to form a O2-H2...O7' hydrogen bond slightly decreases. This accounts for the slightly weaker O2-H2...O7' hydrogen bond strength in the three observed salicin conformers (I, II, and III) compared to V and VIII. It is important to mention here that other higher energy conformers of salicin in the P and Q groups also have  $n \rightarrow \pi^*$  interactions, albeit much weaker in strength compared to those in the observed conformers. This shows that the  $n \rightarrow \pi^*$  interaction plays an important role in the preferential stabilization of the three lowest energy conformers.

Further insights into the nature of the intramolecular interaction between the benzylic OH group and the phenyl ring in salicin are presented in Table 2 as an F/I-SAPT partitioning for conformers I, II, III, V and VIII. The zeroth-order densities of the selected functional groups for conformers I and V are shown as Figure S5 in the SI. It is immediately apparent from the SAPT partitioning that the interaction in I, II and III is roughly the same, with conformers V and VIII forming a group with a different type of interaction. The largest difference between the two groups is that the electrostatic repulsion is significantly weaker for I, II and III, by around a factor of five. The attractive dispersion component is roughly 0.2 kcal mol<sup>-1</sup> stronger for I, II and III, while the induction AB term is on the order of 1 kcal mol<sup>-1</sup> more attractive for the same group of three conformers. Once the SAPT terms are summed it can be seen that the total interaction is repulsive in all cases, but that the interaction between these functional groups is ~5 kcal mol<sup>-1</sup> less repulsive for conformers I, II and III than for V and VIII. The exact quantitative difference here should not be overinterpreted; due to the nature of intramolecular SAPT and the large number of non-bonded interactions in salicin the partitioning captures only a particular slice of the total interaction. However, it is clear that the OH...phenyl interaction is markedly different for the two groups of conformers and that there is a blend of intermolecular forces responsible for the change.

**Table 2:** F/I-SAPT partitioning (kcal mol<sup>-1</sup>) of the interaction between the benzylic OH (functional group A) and phenyl ring (B) in conformers I, II, III, V and VIII of salicin.

SAPT term	Conformer				
	I	II	III	V	VIII
<b>Electrostatic</b>	1.74	1.94	1.81	9.88	9.85
<b>Exchange</b>	8.19	8.05	8.14	4.53	4.53
<b>Induction AB</b>	-2.83	-2.89	-2.83	-1.79	-1.88
<b>Induction BA</b>	-0.30	-0.26	-0.31	-0.82	-0.73
<b>Dispersion</b>	-1.67	-1.68	-1.68	-1.50	-1.49

Previous SAPT calculations on  $n \rightarrow \pi^*$  interactions have focused on intermolecular interactions where the stabilization of the complex was attributed exclusively to said  $n \rightarrow \pi^*$ .<sup>30, 31</sup> In such cases it was found that the electrostatic term was usually repulsive, and that dispersion is the most attractive term. It is worth bearing in mind that different levels and implementations of SAPT do group the individual terms in different ways, often making direct comparisons difficult, even without the added complication of inter- vs. intra-molecular SAPT. For example, the separate  $\delta(\text{HF})$  term from Ref. 30 is included in the induction term in the current investigation (see the PSI4 V1.1 user manual for details of the grouping used in F/I-SAPT), hence it would be expected that induction will make a somewhat greater contribution here when compared to the work of Ran and Hobza.<sup>30</sup> The NBO analysis above indicated that all five conformers considered here possess  $n \rightarrow \pi^*$  interactions, but that the interaction is stronger for conformers I, II and III due to additional donation from  $n_p$ -type electrons. The results of Table 2 are consistent with this, the interactions in all conformers have reasonably strong dispersion components, along with repulsive electrostatic terms. The additional  $n \rightarrow \pi^*$  character in conformers I, II and III then further increases the amount of dispersion and induction, and decreases the magnitude of the electrostatic repulsion.

The effect of dispersion on the equilibrium geometries of salicin conformers has been assessed by comparing structures optimized with the B3LYP and B3LYP-D functionals. Table S10 in the SI shows that inclusion of a dispersion correction leads to a mean average RMSD of 0.045 Å over conformers I-X, indicating that dispersion does have non-negligible effects on the structure of salicin, as would be expected based on the dispersion contributions to  $n \rightarrow \pi^*$  interactions and the large number of intramolecular interactions with salicin. The change in the O7' to phenyl ring centroid distance for B3LYP/B3LYP-D provides some evidence of the contribution of dispersion to the interaction between the benzylic OH and the phenyl ring. Table S10 demonstrates that this distance undergoes a larger change for the group of conformers I, II and III than for V and VIII, indicating an increased role of dispersion in the three lowest energy conformers. However, the difference in the dispersion effect between the two groups of conformers is relatively small at around 0.004 Å. This is entirely consistent with the modest increase in the F/I-SAPT dispersion term in Table 2 that is discussed above.

## Biological perspective

In general, the role of the  $n \rightarrow \pi^*$  interaction in providing conformational preferences of biomolecules and biomolecular complexes is also revealed from protein data bank (PDB) searches.<sup>5, 6, 14, 15, 86, 87</sup> We have found that the preference for the gauche conformation of the isolated benzyl alcohol is retained in the benzylic alcohol moiety of not only isolated salicin but also salicin bound to  $\beta$ -glucosidase enzyme (NkBg1), which is in the glycosyl hydrolase family.<sup>88</sup> This enzyme plays an important role in the cleavage of the  $\beta$ -glucosidic linkage in salicin (a glucose substituted molecule) or any disaccharide. The X-ray crystal structure of the salicin bound complex of  $\beta$ -glucosidase enzyme (PDB ID: 3vil) shows that the benzylic alcohol moiety of salicin adopts a gauche conformation.<sup>88</sup> An NBO analysis of the unoptimized geometry of the salicin moiety obtained from the crystal structure of the salicin... $\beta$ -glucosidase complex carried out at the M06-2X/6-311++G(d,p) level provides an  $n \rightarrow \pi^*$  interaction energy of 1.04 kcal/mol (see SI), which is similar to that presented in Table 1 for isolated salicin. The PDB structure highlighting the binding pocket of the enzyme with salicin and NBOs of the salicin moiety are presented in Figure S4 of the SI.

The propensity of the  $n \rightarrow \pi^*$  interaction and hydrogen bonds to both stabilize and influence the shape of protein structures has been pointed out by Raines and coworkers as a result of analyzing the crystal structure of proteins in the PDB.<sup>5, 6</sup> Hydrogen bonding ( $n \rightarrow \sigma^*$ ) and  $n \rightarrow \pi^*$  interactions are interrelated as both interactions originate from the delocalization of lone pair electrons. In proteins, mostly in  $\alpha$ -helices, lone pair electrons of the same oxygen atom of a C=O group are often shared by both hydrogen bonding and an  $n \rightarrow \pi^*$  interaction. For example, there is a 50:50 sharing of the p-type lone pair electrons of the carbonyl oxygen atom by the hydrogen bond and the  $n \rightarrow \pi^*$  interaction in the asparagine residues of a human carbonic anhydrase-II protein (PDB 3KS3).<sup>6</sup> Similar results were obtained for salicin in this work. In the case of the three observed conformers of salicin, p-type lone pair electrons on the oxygen atom of the benzylic alcohol moiety participate in both O...H-O hydrogen bonding and the  $n \rightarrow \pi^*$  interaction. It is also observed that the strength of the hydrogen bond in the three observed conformers of salicin decreases due to the presence of a significant amount of  $n \rightarrow \pi^*$  interaction compared to that in the other conformers of salicin.

## Conclusion

Conformational preferences of salicin induced by a subtle interplay between the  $n \rightarrow \pi^*$  interactions and hydrogen bond have been studied using isolated gas phase electronic and vibrational spectroscopy combined with quantum chemistry calculations. Three low energy conformers of salicin, which are observed in the experiment, have a significant amount of  $n \rightarrow \pi^*$  interaction along with the hydrogen bonding interactions when compared to other higher energy conformers. The results of NBO, QTAIM and SAPT calculations indicate that the higher energy conformers have comparable hydrogen bonding interactions, but weaker  $n \rightarrow \pi^*$  interactions. Sharing of lone pair electrons on a single oxygen atom between hydrogen bonding and  $n \rightarrow \pi^*$  interactions reported for the stabilization of protein structures has also been observed in salicin here. Thus, the present research suggests that the presence of a hydrogen bonding interaction in a molecular system can very often indicate the possibility of the existence of a  $n \rightarrow \pi^*$  interaction as these two non-covalent interactions are quite analogous in terms of electron delocalization. This indicates that the  $n \rightarrow \pi^*$  interaction should be incorporated into the simulation of biomolecules to obtain more accurate results for their structures and dynamics.

The present work also adds further evidence of competition and cooperation between weak intermolecular interactions in determining the low-energy conformers of floppy drug molecules. In this case, a slight compromise in the strength of the hydrogen bonding interaction favors the  $n \rightarrow \pi^*$  interaction in the overall stability of the molecular structures. Another intriguing finding of this study is that the conformational preference of the gauche conformer of either isolated benzyl alcohol or the benzyl alcohol moiety of salicin over the trans conformer is not due to the O-H... $\pi$  interaction, but rather an  $n \rightarrow \pi^*$  interaction. Therefore, it is proposed that the  $n \rightarrow \pi^*$  interaction could have a significant contribution in the stabilization of molecular systems that have previously been reported to be stabilized by O-H... $\pi$  interactions. The contribution of the  $n \rightarrow \pi^*$  interaction in O-H... $\pi$  bound systems has largely been overlooked in previous studies. In future, it will be interesting to reinvestigate systems containing O-H... $\pi$  interactions to probe for the presence of an  $n \rightarrow \pi^*$  interaction there.

## Notes

The authors declare no competing financial interest.

## Acknowledgments

SKS, PRJ, and AD would like to acknowledge IISER Pune for providing the financial assistance to perform this research. Computational support from IISER Pune HPC cluster is acknowledged. SKS thanks IISER Pune for providing a research fellowship. Partial financial support received from DST-SERB (EMR/2015/000486) to execute this research is gratefully acknowledged.

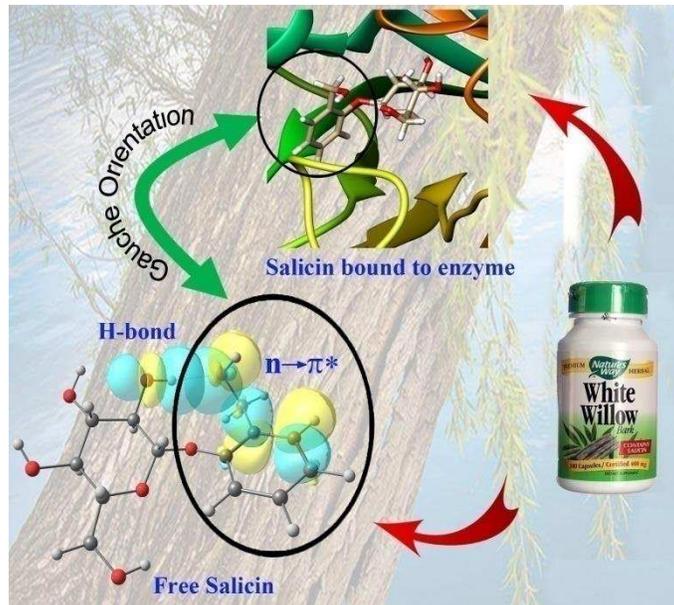
## References

1. C. B. Anfinsen, *Science*, 1973, **181**, 223-230.
2. A. Karshikoff, *Non-covalent Interactions in Proteins*, Imperial College Press, 2006.
3. R. W. Newberry and R. T. Raines, *Acc. Chem. Res.*, 2017, **50**, 1838-1846.
4. S. K. Singh and A. Das, *Phys. Chem. Chem. Phys.*, 2015, **17**, 9596-9612.
5. G. J. Bartlett, A. Choudhary, R. T. Raines and D. N. Woolfson, *Nat. Chem. Biol.*, 2010, **6**, 615-620.
6. G. J. Bartlett, R. W. Newberry, B. VanVeller, R. T. Raines and D. N. Woolfson, *J. Am. Chem. Soc.*, 2013, **135**, 18682-18688.
7. R. W. Newberry, B. VanVeller, I. A. Guzei and R. T. Raines, *J. Am. Chem. Soc.*, 2013, **135**, 7843-7846.
8. J. A. Hodges and R. T. Raines, *Org. Lett.*, 2006, **8**, 4695-4697.
9. R. W. Newberry, G. J. Bartlett, B. VanVeller, D. N. Woolfson and R. T. Raines, *Protein Sci.*, 2014, **23**, 284-288.
10. M. P. Hinderaker and R. T. Raines, *Protein Sci.*, 2003, **12**, 1188-1194.
11. S. K. Singh, S. Kumar and A. Das, *Phys. Chem. Chem. Phys.*, 2014, **16**, 8819-8827.
12. S. K. Singh, J. K. Vaishnav and A. Das, *J. Chem. Phys.*, 2016, **145**, 104302.
13. S. K. Singh, A. Das and G. W. Breton, *J. Phys. Chem. A*, 2016, **120**, 6258-6269.
14. M. Egli and S. Sarkhel, *Acc. Chem. Res.*, 2007, **40**, 197-205.
15. M. Egli and R. V. Gessner, *Proc. Natl. Acad. Sci. U.S.A.*, 1995, **92**, 180-184.
16. A. Choudhary, D. Gandla, G. R. Krow and R. T. Raines, *J. Am. Chem. Soc.*, 2009, **131**, 7244-7246.
17. B. C. Gorske, B. L. Bastian, G. D. Geske and H. E. Blackwell, *J. Am. Chem. Soc.*, 2007, **129**, 8928-8929.
18. B. C. Gorske, J. R. Stringer, B. L. Bastian, S. A. Fowler and H. E. Blackwell, *J. Am. Chem. Soc.*, 2009, **131**, 16555-16567.
19. T. K. Pal and R. Sankararamkrishnan, *J. Phys. Chem. B*, 2010, **114**, 1038-1049.
20. S. Blanco, J. C. López, S. Mata and J. L. Alonso, *Angew. Chem. Int. Ed.*, 2010, **49**, 9187-9192.
21. A. Choudhary, K. J. Kamer and R. T. Raines, *J. Org. Chem.*, 2011, **76**, 7933-7937.
22. T. J. Mooibroek, P. Gamez and J. Reedijk, *CrystEngComm*, 2008, **10**, 1501-1515.
23. A. J. Neel, M. J. Hilton, M. S. Sigman and F. D. Toste, *Nature*, 2017, **543**, 637-646.
24. M. L. DeRider, S. J. Wilkens, M. J. Waddell, L. E. Bretscher, F. Weinhold, R. T. Raines and J. L. Markley, *J. Am. Chem. Soc.*, 2002, **124**, 2497-2505.
25. R. W. Newberry and R. T. Raines, *Chem Commun*, 2013, **49**, 7699-7701.

26. R. W. Newberry and R. T. Raines, *Acs Chem. Biol.*, 2014, **9**, 880-883.
27. J. Kozelka, *Eur. Biophys. J.*, 2017, **46**, 729-737.
28. J. Novotny, S. Bazzi, R. Marek and J. Kozelka, *Phys. Chem. Chem. Phys.*, 2016, **18**, 19472-19481.
29. Z. Badri, C. Foroutan-Nejad, J. Kozelka and R. Marek, *Phys. Chem. Chem. Phys.*, 2015, **17**, 26183-26190.
30. J. Ran and P. Hobza, *J. Chem. Theory Comput.*, 2009, **5**, 1180-1185.
31. T. Yang, J. J. An, X. Wang, D. Y. Wu, W. Chen and J. S. Fossey, *Phys. Chem. Chem. Phys.*, 2012, **14**, 10747-10753.
32. S. K. Singh, K. K. Mishra, N. Sharma and A. Das, *Angew. Chem. Int. Ed.*, 2016, **55**, 7801-7805.
33. A. Rahim, P. Saha, K. K. Jha, N. Sukumar and B. K. Sarma, *Nat Commun*, 2017, **8**, doi: 10.1038/s41467-017-00081-x.
34. M. Shara and S. J. Stohs, *Phytother. Res.*, 2015, **29**, 1112-1116.
35. J. Vlachojannis, F. Magora and S. Chrubasik, *Phytother. Res.*, 2011, **25**, 1102-1104.
36. G. Ali, F. Subhan, A. Wadood, A. Khan, N. Ullah, N. U. Islam and I. Khan, *Afr. J. Pharm. Pharmacol.*, 2013, **7**, 585-596.
37. J. G. Mahdi, *J. Saudi Chem. Soc.*, 2010, **14**, 317-322.
38. J. G. Mahdi, *Biotechnol. Rep.*, 2014, **4**, 73-79.
39. R. Julkunen-Tiitto and B. Meier, *J. Nat. Prod.*, 1992, **55**, 1204-1212.
40. C.-S. Kong, K.-H. Kim, J.-S. Choi, J.-E. Kim, C. Park and J.-W. Jeong, *Phytother. Res.*, 2014, **28**, 1246-1251.
41. U. Wölfle, B. Haarhaus, A. Kersten, B. Fiebich, M. J. Hug and C. M. Schempp, *Phytother. Res.*, 2015, **29**, 1494-1500.
42. K. Hostanska, G. Jürgenliemk, G. Abel, A. Nahrstedt and R. Saller, *Cancer Detect Prev.*, 2007, **31**, 129-139.
43. S. Preeti, V. K. Singh, S. Brahma Deo, S. Gaurava, B. B. Misra and T. Vyasji, *J. Proteomics Bioinform.* **2013**, **6**, 109-124.
44. G. A. Bonaterra, E. U. Heinrich, O. Kelber, D. Weiser, J. Metz and R. Kinscherf, *Phytomedicine*, 2010, **17**, 1106-1113.
45. S. Kumar, S. K. Singh, C. Calabrese, A. Maris, S. Melandri and A. Das, *Phys. Chem. Chem. Phys.*, 2014, **16**, 17163-17171.
46. C. E. H. Dessent, W. D. Geppert, S. Ullrich and K. M. Dethlefs, *Chem. Phys. Lett.*, 2000, **319**, 375-384.
47. F. O. Talbot and J. P. Simons, *Phys. Chem. Chem. Phys.*, 2002, **4**, 3562-3565.
48. R. A. Jockusch, R. T. Kroemer, F. O. Talbot and J. P. Simons, *J. Phys. Chem. A*, 2003, **107**, 10725-10732.
49. I. Usabiaga, J. Gonzalez, P. F. Arnaiz, I. Leon, E. J. Cocinero and J. A. Fernandez, *Phys. Chem. Chem. Phys.*, 2016, **18**, 12457-12465.
50. S. Kumar, P. Biswas, I. Kaul and A. Das, *J. Phys. Chem. A*, 2011, **115**, 7461.
51. S. Kumar, I. Kaul, P. Biswas and A. Das, *J. Phys. Chem. A*, 2011, **115**, 10299-10308.
52. S. Kumar and A. Das, *J. Chem. Phys.*, 2012, **136**, 174302-174309.
53. S. Kumar, A. Mukherjee and A. Das, *J. Phys. Chem. A*, 2012, **116**, 11573-11580.
54. G. Meijer, M. S. de Vries, H. E. Hunziker and H. R. Wendt, *Applied Physics B*, 1990, **51**, 395-403.
55. F. Piuze, I. Dimicoli, M. Mons, B. Tardivel and Q. Zhao, *Chem. Phys. Lett.*, 2000, **320**, 282-288.
56. H. Goto and E. Osawa, *J. Am. Chem. Soc.*, 1989, **111**, 8950-8951.
57. H. Goto and E. Osawa, *J. Chem. Soc., Perkin Trans. 2*, 1993, 187-198.
58. H. Goto and E. Osawa, CONFLEX 3 evaluation version, Japan Chemistry Program Exchange (JCPE), P40.
59. Y. Zhao and D. G. Truhlar, *Theor. Chem. Acc.*, 2008, **120**, 215-241.

60. N. Mardirossian and M. Head-Gordon, *J. Chem. Theory Comput.*, 2016, **12**, 4303-4325.
61. M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, et al. Gaussian 09, revision D.01; Gaussian Inc.: Wallingford, CT, 2009.
62. A. D. Laurent and D. Jacquemin, *Int. J. Quantum Chem.*, 2013, **113**, 2019-2039.
63. D. Jacquemin, E. A. Perpète, I. Ciofini and C. Adamo, *J. Chem. Theory Comput.*, 2010, **6**, 1532-1537.
64. V. Barone, J. Bloino, M. Biczysko and F. Santoro, *J. Chem. Theory Comput.*, 2009, **5**, 540-554.
65. Schriever, C.; Cockett, M. C. R.; Pugliesi, I. FC-LabII, Version 2009a. The latest information on program updates, a basic introduction to Franck-Condon simulations and a free download of the software can be found at <http://www.fclab2.net/>.
66. [http://www.nist.gov/mm1/csd/informatics\\_research/thermochemistry\\_script.cfm](http://www.nist.gov/mm1/csd/informatics_research/thermochemistry_script.cfm)).
67. F. Weinhold and C. R. Landis, *Valency and Bonding: A Natural Bond Orbital Donor-Acceptor Perspective*, Cambridge University Press, 2005.
68. E. D. Glendening, C. R. Landis and F. Weinhold, *J. Comput. Chem.*, 2013, **34**, 1429-1437.
69. B. Jeziorski, R. Moszynski and K. Szalewicz, *Chem. Rev.*, 1994, **94**, 1887-1930.
70. R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince, E. G. Hohenstein, U. Bozkaya, A. Y. Sokolov, R. Di Remigio, R. M. Richard, et al. *J. Chem. Theory Comput.*, 2017, **13**, 3185-3197.
71. T. H. D. Jr., *J. Chem. Phys.*, 1989, **90**, 1007-1023.
72. R. A. Kendall, T. H. D. Jr. and R. J. Harrison, *J. Chem. Phys.*, 1992, **96**, 6796-6806.
73. R. M. Parrish, T. M. Parker and C. D. Sherrill, *J. Chem. Theory Comput.*, 2014, **10**, 4417-4431.
74. R. M. Parrish, J. F. Gonthier, C. Corminbœuf and C. D. Sherrill, *J. Chem. Phys.*, 2015, **143**, 051103.
75. A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648-5652.
76. P. J. Stephens, F. J. Devlin, C. F. Chabalowski and M. J. Frisch, *J. Phys. Chem.*, 1994, **98**, 11623-11627.
77. G. Stefan, E. Stephan and G. Lars, *J. Comput. Chem.*, 2011, **32**, 1456-1465.
78. M. W. Walker, L. Shao and R. A. Volz, *CVGIP: Image Understanding*, 1991, **54**, 358-367.
79. G. Calculate Root-mean-square deviation (RMSD) of Two Molecules Using Rotation, <http://github.com/charnley/rmsd>, version 1.2.7.
80. S. Coussan, Y. Bouteiller, J. P. Perchard and W. Q. Zheng, *J. Phys. Chem. A*, 1998, **102**, 5789-5793.
81. M. Mons, E. G. Robertson and J. P. Simons, *J. Phys. Chem. A*, 2000, **104**, 1430-1437.
82. S. W. Han and K. Kim, *J. Phys. Chem.*, 1996, **100**, 17124-17132.
83. M. P. Gosling, I. Pugliesi and M. C. R. Cockett, *Phys. Chem. Chem. Phys.*, 2010, **12**, 132-142.
84. J. G. Hill and A. Das, *Phys. Chem. Chem. Phys.*, 2014, 2014, **16**, 11754-11762.
85. E. Arunan, G. R. Desiraju, R. A. Klein, J. Sadlej, S. Scheiner, I. Alkorta, D. C. Clary, R. H. Crabtree, J. J. Dannenber, P. Hobza, H. G. Kjaergaard, A. C. Legon, B. Mennucci and D. J. Nesbitt, *Pure Appl. Chem.*, 2011, **83**, 1637-1641.
86. A. Jain, V. Ramanathan and R. Sankararamakrishnan, *Protein Sci.*, 2009, **18**, 595-605.
87. A. Jain, C. S. Purohit, S. Verma and R. Sankararamakrishnan, *J. Phys. Chem. B*, 2007, **111**, 8680-8683.
88. W.-Y. Jeng, N.-C. Wang, C.-T. Lin, W.-J. Chang, C.-I. Liu and A. H. J. Wang, *Acta Crystallogr., Sect. D*, 2012, **68**, 829-838.

## Table of content



$n \rightarrow \pi^*$  interaction is present in the structure of salicin when it is bound to enzyme as well as in free state and the conformational preference of salicin is due to interplay between strong hydrogen bond and  $n \rightarrow \pi^*$  interaction.