Epigallocatechin gallate (EGCG) with potent anti-*Helicobacter pylori* activity binds efficiently to its Histone like DNA binding protein

Ritu Raj¹,², Nipanshu Agarwal³, Sriram Raghavan⁴, Tapati Chakraborti², Krishna Mohan Poluri³, Gaurav Pande* and Dinesh Kumar¹,*

¹Centre of Biomedical Research (CBMR), SGPGIMS Campus, Raibareli Road, Lucknow-226014, Uttar Pradesh, India
²Department of Biochemistry & Biophysics, University of Kalyani, Kalyani-741235, West Bengal, India
³Department of Biotechnology and Centre for Nanotechnology Indian Institute of Technology Roorkee
Roorkee – 247667, Uttarakhand, India
⁴Computational Structural Biology Team, RIKEN Center for Computational Science (R-CCS) Kobe-650-0047, Japan
⁵Department of Gastroenterology, SGPGIMS, Lucknow-226014, Uttar Pradesh, India

*Author for Correspondence:

| Dr. Gaurav Pande | Dr. Dinesh Kumar |
|------------------|------------------|
| (DM, MBBS)       | (Associate Professor) |
| Department of Gastroenterology, SGPGIMS, Raibareli Road, Lucknow-226014, Uttar Pradesh, India | Centre of Biomedical Research (CBMR), SGPGIMS Campus, Raibareli Road, Lucknow-226014, Uttar Pradesh, India |
| Email: drgauravpandey@yahoo.com | Email: dineshcbmr@gmail.com |
| Contact: +91-9170689999 | Contact: +91-8953261506 |

Keywords

*Helicobacter pylori*, Histone like DNA Binding Protein (Hup); Proton NMR; Molecular docking; Virtual screening; Molecular Dynamics simulation.
**Figure S1:** Geometry optimization performed for natural compound library containing anti-*Helicobacter pylori* (AHP) compounds (downloaded from the Pubchem database: https://pubchem.ncbi.nlm.nih.gov) using LigPrep(v2.3) application of Schrodinger \(^1\) so that the Pka and protonation state of the compounds can be used to determine possible tautomerization in OPLS3 force field. The input library file contains 159 distinctive chemical compounds and the output file contains 246 conformations corresponding to 133 compounds with molecular weight less than 500 Da. The bond orders were optimised and ionization state of the ligands were generated with optimal pH of 7.0 using Epik(2.2). The display shown below has been derived from Schrodiner’s LigPrep application module [LigPrep, Schrödinger, LLC, New York, NY, 2020] as a reference for future studies.
Figure S2: (A,B) The energy minimization of dimeric Hup model generated using YASARA homology modelling application. Energy minimization is performed using YASARA free webserver (http://www.yasara.org/minimizationserver.htm). The 94 amino acid sequence of \textit{H. pylori} Hup derived from here: https://www.ncbi.nlm.nih.gov/protein/WP_001029082.1 (in FASTA format) is used as input for homology modeling (MNKAEFIDLV KEAGKYNKR EAEAAISAFT LAVETALSKG ESVELIGFGK FETAEQKGKE GKVPGSDKTY KTEDKRVPKF KPGKTLQKV EEGK). (C) The resulted dimeric Hup model was compared to those generated previously\textsuperscript{2} using free web-based Protein Structure prediction server named (PS)\textsuperscript{2}v2.\textsuperscript{3,4} (D) The comparison has been made between energy minimized structures.
Figure S3: The energy minimization of monomeric Hup model performed using YASARA energy minimization webserver. The monomeric Hup structure has been extracted from Chain-B of homodimeric model generated using YASARA homology modelling.
Appendix-I: Molecular Dynamics Simulations

Prior to simulation, the complex was cleaned and optimized for hydrogen bond network. Next, a cubic simulation cell was created around the macromolecular system filled with water molecules and the periodic boundary condition was incorporated to perform the simulation. For charge neutrality as well as to mimic the biological conditions partly, the sodium and chloride (Na⁺/Cl⁻) ions were also added randomly to the system (by replacing the water molecules). The AMBER14 force field was employed to describe the macromolecular atoms in the simulation system. The pKa values of titratable amino acids (e.g. Asp, Glu, His and Lys) were estimated and simulation box was solvated using the transferable intermolecular potential3 points (TIP3P) water model (density: 0.997 g/L) accordingly. Before MD simulation, each macromolecular system was energy minimized employing steepest descent approach (5000 cycles) followed by simulated annealing approach. All-atom MD simulations were performed employing particle-mesh Ewald (PME) method to describe long-range electrostatic interactions at a cut off distance of 8 Å and defining near physiological conditions at 298 K, pH 7.4 and 0.9% NaCl. The simulation temperature was controlled using the Berendsen thermostat, where the pressure kept constant throughout the simulation. A multiple time step algorithm together with a simulation time step interval of 2.50 fs was chosen. Long run MD simulations (more than 100 ns in each case) were performed at constant temperature using a Berendsen thermostat and constant pressure. MD trajectories were saved every 250 ps for further analysis. The MD trajectory data collected was analyzed using YASARA built in macro named “mdanalysis.mcr”. The average structures were determined from the simulations and were used to calculate the Root-mean-square deviation (RMSD), root-mean-square fluctuation (RMSF), and number of H-bonds between solute and solvent for the respective macromolecular system. YASARA has a built-in protocol for calculating binding energies (BindEnergyObj command). The binding free energy (ΔG\text{binding}, referred here as binding energy) is obtained by calculating the energy at infinite distance between the selected object and the rest of the simulation system (i.e., the unbound state) and subtracting the energy of the simulation system (i.e., the bound state). As per the details provided in the YASARA manual, the free energy of an object is calculated without involving entropy term from normal mode analysis as:

\begin{equation}
G = E_{\text{bind}} + E_{\text{el}} + E_{\text{vdw}} + G_{\text{polar}} + G_{\text{nonpolar}} - TS \quad (S1)
\end{equation}

Here, the first three terms are molecular mechanics (MM) terms of binding, electrostatic, and van der Waals interactions, respectively. Gpolar and Gnonpolar are polar and nonpolar contributions of solvation free energies, respectively. TS is the entropy effect, which is neglected when binding affinities are compared for the same ligand in different binding sites. Then binding free energy is calculated by the following equation:

\begin{equation}
\Delta G_{\text{binding}} = \left( G_{\text{Receptor internal}} + G_{\text{Ligand internal}} \right) + \left( G_{\text{Receptor solvation}} + G_{\text{Ligand solvation}} \right) - \left( G_{\text{complex internal}} + G_{\text{complex solvation}} \right) \quad [ \text{kJ/mol} ] \quad (S2)
\end{equation}
Here, the first two terms are potential energies of the receptor and ligand, the next two terms are solvation energies of the receptor and ligand, and the last terms are potential and solvation energies of the complex. As mentioned in the help manual, YASARA sticks to standard convention and provides positive binding energies in KJ/mol. Note that more positive the binding energy, the more favorable the interaction in the context of the chosen force field, whereas negative energies indicates very weak binding, but DO NOT indicate no binding. Depending upon the system, more negative binding energy values indicate more favorable binding. For example, the negative binding energies (calculated using YASARA) has been reported previously for ligands binding to membrane-bound protein. Relevant binding contacts between receptor and ligand were detected and visualized making composite use of YASARA software applications.

Table S1: The binding energy parameters derived from virtual screening of selected compound library against dimeric and monomeric Hup structures employing small molecule docking approaches (AUTODOCK, VINA and Glide). Natural compounds with highest cumulative binding energy were highlighted (with * mark and in Red color) for convenience.

| ID | Autodock BE (Kcal/mol) | VINA BE (Kcal/mol) | Glide BE (Kcal/mol) | CBE (Kcal/mol) | Pubmed ID (CID) |
|----|------------------------|--------------------|---------------------|----------------|-----------------|
|    | HupM | HupD | HupM | HupD | HupM | HupD |                     |                |
| 1  | 6.93 | 7.32 | 6.74 | 6.58 | 3.50 | 0.58 | 5.27 | 2353               |
| 2  | 5.84 | 6.95 | 6.44 | 5.53 | 4.74 | 4.31 | 5.64 | 72300              |
| 3  | 6.35 | 7.56 | 6.58 | 5.51 | 4.42 | 0.00 | 5.07 | 72300              |
| 4  | 6.08 | 6.61 | 7.85 | 6.31 | 3.64 | 3.00 | 5.58 | 442126             |
| 5  | 5.15 | 6.04 | 4.66 | 5.05 | 4.49 | 5.15 | 5.09 | 370                |
| 6  | 5.75 | 6.17 | 5.10 | 5.29 | 2.38 | 2.14 | 4.47 | 444539             |
| 7  | 6.31 | 7.58 | 6.32 | 6.16 | 4.55 | 2.42 | 5.56 | 25201019           |
| 8  | 6.18 | 7.03 | 6.21 | 6.21 | 4.40 | 5.16 | 5.87 | 25201019           |
| 9  | 6.19 | 7.71 | 6.98 | 6.47 | 4.13 | 5.72 | 6.20 | 5281789            |
| 10 | 6.85 | 7.81 | 7.03 | 6.40 | 4.14 | 5.30 | 6.26 | 5281789            |
| 11 | 6.42 | 7.42 | 5.77 | 5.61 | 4.34 | 4.56 | 5.69 | 5281781            |
| 12 | 6.83 | 8.17 | 5.94 | 5.78 | 3.59 | 5.78 | 6.01 | 5281781            |
| 13 | 6.38 | 7.45 | 6.01 | 5.82 | 4.62 | 2.10 | 5.40 | 5280961            |
| 14 | 6.87 | 7.80 | 6.17 | 5.75 | 4.75 | 5.00 | 6.06 | 5280961            |
| 15 | 6.40 | 7.08 | 6.12 | 6.13 | 3.39 | 2.32 | 5.24 | 628528             |
| 16 | 6.37 | 7.24 | 6.23 | 6.24 | 4.22 | 3.44 | 5.62 | 72281              |
| 17 | 5.95 | 6.48 | 4.96 | 5.15 | 2.61 | 1.65 | 4.47 | 637542             |
| 18 | 6.54 | 6.89 | 4.99 | 5.14 | 3.54 | 4.27 | 5.23 | 689043             |
| 19 | 8.46 | 9.09 | 6.60 | 6.33 | 5.65 | 3.75 | 6.65 | 5281792            |
| 20 | 5.51 | 5.93 | 5.40 | 5.46 | 3.45 | 4.34 | 5.02 | 11644379           |
| 21 | 5.03 | 5.59 | 5.55 | 5.56 | 3.91 | 2.92 | 4.76 | 11218565           |
|     |       |       |       |       |       |       |       |       |       |       |
|-----|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 22  | 5.32  | 5.99  | 5.37  | 4.93  | 2.79  | 1.56  | 4.33  | 11218565 |
| 23  | 6.00  | 6.26  | 5.60  | 5.44  | 4.18  | 0.00  | 4.58  | 6857681  |
| 24  | 5.32  | 6.35  | 5.43  | 4.55  | 2.95  | 0.00  | 4.10  | 15560069 |
| 25  | 6.98  | 8.75  | 8.91  | 8.04  | 3.49  | 5.12  | 6.88  | 10114   |
| 26  | 5.01  | 5.11  | 4.04  | 4.40  | 3.40  | 5.04  | 4.50  | 1183    |
| 27  | 5.23  | 5.63  | 4.31  | 4.19  | 2.82  | 3.69  | 4.31  | 1183    |
| 28  | 6.72  | 7.51  | 7.10  | 6.03  | 4.22  | 4.22  | 5.97  | 14287147|
| 29  | 6.54  | 7.76  | 7.18  | 6.56  | 3.29  | 3.48  | 5.80  | 14680349|
| 30  | 7.44  | 7.78  | 7.33  | 5.94  | 4.40  | 2.40  | 5.88  | 101297697|
| 31  | 5.72  | 6.61  | 5.15  | 5.35  | 3.96  | 2.83  | 4.94  | 44584263|
| 32  | 6.32  | 7.17  | 5.59  | 5.63  | 3.96  | 0.28  | 4.82  | 44584263|
| 33  | 6.43  | 7.39  | 6.06  | 5.94  | 3.48  | 3.93  | 5.54  | 5352001 |
| 34  | 7.19  | 7.59  | 6.06  | 6.05  | 3.79  | 5.26  | 5.99  | 5352001 |
| 35  | 6.32  | 6.83  | 6.04  | 6.12  | 4.58  | 2.53  | 5.40  | 5459196 |
| 36  | 7.03  | 7.58  | 5.95  | 6.27  | 4.01  | 2.72  | 5.59  | 5459196 |
| 37  | 6.17  | 7.30  | 6.01  | 6.08  | 3.91  | 5.02  | 5.75  | 5281666 |
| 38  | 6.97  | 7.77  | 6.07  | 6.29  | 4.84  | 6.94  | 6.48  | 5281666 |
| 39  | 4.92  | 5.57  | 4.95  | 4.68  | 3.16  | 2.03  | 4.22  | 7444    |
| 40  | 4.71  | 6.98  | 5.11  | 4.23  | 4.11  | 0.00  | 4.19  | 167551  |
| 41  | 6.50  | 7.89  | 6.25  | 6.05  | 3.81  | 3.01  | 5.58  | 443023  |
| 42  | 5.67  | 5.72  | 4.92  | 5.14  | 3.67  | 2.48  | 4.60  | 3806    |
| 43  | 6.28  | 6.52  | 4.93  | 5.06  | 2.77  | 0.26  | 4.30  | 3806    |
| 44  | 6.59  | 7.90  | 6.07  | 5.86  | 3.91  | 3.25  | 5.60  | 92503   |
| 45  | 6.54  | 7.78  | 6.03  | 5.98  | 3.12  | 6.63  | 6.01  | 92503   |
| 46  | 6.01  | 7.02  | 6.37  | 6.00  | 3.85  | 7.64  | 6.15  | 5319013 |
| 47  | 6.23  | 7.88  | 6.46  | 6.20  | 4.01  | 7.37  | 6.36  | 5319013 |
| 48  | 6.35  | 6.31  | 6.82  | 6.41  | 4.84  | 2.60  | 5.56  | 480873  |
| 49  | 6.53  | 7.62  | 6.90  | 6.69  | 3.07  | 3.33  | 5.69  | 480865  |
| 50  | 6.37  | 7.32  | 6.56  | 6.16  | 4.08  | 3.24  | 5.62  | 114829  |
| 51  | 6.55  | 8.05  | 7.19  | 6.70  | 4.23  | 5.79  | 6.42  | 124052  |
| 52  | 7.02  | 7.62  | 7.26  | 7.18  | 4.60  | 3.84  | 6.25  | 480774  |
| 53  | 6.79  | 7.27  | 6.82  | 6.44  | 3.38  | 4.78  | 5.91  | 5320083 |
| 54  | 7.43  | 8.29  | 6.96  | 6.75  | 3.74  | 4.76  | 6.32  | 5320083 |
| 55  | 6.97  | 9.14  | 6.29  | 6.06  | 5.72  | 1.42  | 5.93  | 1794427 |
| 56  | 6.79  | 7.37  | 6.20  | 6.21  | 3.79  | 2.95  | 5.55  | 440735  |
| 57  | 6.74  | 7.47  | 6.06  | 6.28  | 5.22  | 3.30  | 5.84  | 439533  |
| 58  | 7.38  | 8.48  | 6.01  | 6.42  | 4.46  | 4.64  | 6.23  | 439533  |
| 59  | 6.60  | 7.87  | 5.89  | 6.18  | 6.01  | 3.00  | 6.32  | 5320083 |
| 60  | 7.44  | 8.48  | 6.10  | 6.29  | 4.76  | 4.91  | 6.33  | 5281691 |
| 61  | 6.35  | 6.72  | 6.06  | 5.96  | 4.18  | 1.89  | 5.19  | 6442633 |
| 62  | 6.06  | 6.06  | 6.15  | 5.88  | 4.13  | 3.18  | 5.24  | 6442633 |
| 63  | 6.90  | 6.89  | 6.25  | 6.01  | 4.10  | 2.71  | 5.48  | 6442633 |
|   | Column 1 | Column 2 | Column 3 | Column 4 | Column 5 | Column 6 |
|---|----------|----------|----------|----------|----------|----------|
| 64| 6.31     | 6.55     | 6.24     | 5.74     | 3.17     | 3.99     |
| 65| 6.78     | 7.41     | 6.16     | 5.80     | 4.09     | 3.88     |
| 66| 6.39     | 6.53     | 6.11     | 6.00     | 4.01     | 3.80     |
| 67| 6.66     | 7.28     | 5.95     | 6.13     | 3.99     | 3.79     |
| 68| 6.47     | 7.00     | 5.93     | 5.80     | 3.98     | 3.62     |
| 69| 6.82     | 7.06     | 6.14     | 5.67     | 3.97     | 3.54     |
| 70| 6.30     | 6.96     | 6.06     | 5.68     | 3.93     | 0.79     |
| 71| 5.94     | 6.67     | 6.37     | 5.98     | 3.92     | 3.48     |
| 72| 6.40     | 6.82     | 6.15     | 6.64     | 3.91     | 3.46     |
| 73| 7.05     | 6.86     | 6.00     | 5.93     | 3.91     | 3.44     |
| 74| 6.32     | 6.86     | 6.05     | 5.84     | 3.88     | 3.39     |
| 75| 7.06     | 7.14     | 6.38     | 6.32     | 3.85     | 3.36     |
| 76| 6.94     | 6.86     | 6.02     | 6.39     | 3.84     | 3.36     |
| 77| 6.13     | 6.29     | 6.15     | 5.79     | 3.83     | 3.28     |
| 78| 6.28     | 7.16     | 6.44     | 6.19     | 3.82     | 3.22     |
| 79| 7.04     | 6.93     | 6.15     | 6.21     | 3.80     | 3.20     |
| 80| 6.70     | 6.39     | 6.09     | 6.35     | 3.78     | 4.25     |
| 81| 6.86     | 6.99     | 5.99     | 6.02     | 3.78     | 3.18     |
| 82| 6.29     | 7.05     | 6.22     | 6.19     | 3.70     | 3.17     |
| 83| 6.50     | 7.52     | 6.14     | 5.87     | 3.58     | 3.16     |
| 84| 6.66     | 7.61     | 6.11     | 6.23     | 3.58     | 3.07     |
| 85| 6.39     | 7.36     | 6.29     | 5.97     | 3.57     | 2.97     |
| 86| 6.39     | 6.20     | 6.32     | 5.70     | 3.55     | 2.94     |
| 87| 5.90     | 7.04     | 5.83     | 6.09     | 3.55     | 2.90     |
| 88| 6.81     | 6.83     | 6.59     | 6.34     | 3.54     | 2.87     |
| 89| 6.33     | 6.82     | 6.02     | 5.96     | 3.53     | 2.72     |
| 90| 6.46     | 6.56     | 6.31     | 5.85     | 3.49     | 4.12     |
| 91| 6.52     | 6.50     | 6.41     | 6.20     | 3.40     | 2.65     |
| 92| 7.09     | 6.78     | 6.05     | 6.03     | 3.33     | 1.54     |
| 93| 7.32     | 6.92     | 6.37     | 6.04     | 3.40     | 7.51     |
| 94| 5.87     | 7.05     | 6.05     | 5.78     | 4.10     | 5.15     |
| 95| 7.37     | 7.87     | 6.28     | 6.76     | 5.38     | -3.00    |
| 96*| 8.47     | 8.91     | 6.43     | 6.63     | 5.46     | 8.33     |
| 97| 8.33     | 8.99     | 6.39     | 7.24     | 5.81     | 7.28     |
| 98| 4.96     | 6.24     | 5.43     | 5.07     | 2.40     | 0.00     |
| 99| 7.16     | 7.80     | 6.16     | 6.20     | 5.28     | 3.29     |
|100| 8.14     | 8.79     | 6.32     | 6.38     | 4.89     | 5.81     |
|101| 8.20     | 8.85     | 6.39     | 6.38     | 4.85     | 2.49     |
|102| 6.73     | 7.83     | 6.07     | 6.09     | 5.55     | 5.11     |
|103| 7.51     | 8.75     | 5.92     | 6.24     | 5.29     | 4.79     |
|104| 6.58     | 7.72     | 6.93     | 6.98     | 4.36     | 10.02    |
|105| 6.37     | 7.93     | 6.88     | 6.85     | 5.33     | 3.80     |
|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
|106| 6.56| 8.15| 6.67| 6.67| 5.14| 3.76|
|107| 6.43| 8.17| 6.70| 6.40| 4.78| 2.32|
|108| 6.50| 7.92| 6.70| 6.33| 4.03| 0.93|
|109| 4.37| 4.65| 5.41| 4.48| 3.20| 5.34|
|110| 4.83| 5.16| 4.21| 4.16| 2.50| 3.46|
|111| 4.17| 4.39| 3.79| 4.65| 3.70| 2.38|
|112| 3.78| 4.22| 3.35| 3.25| 2.16| 1.48|
|113| 3.66| 4.14| 3.22| 3.40| 2.11| 4.50|
|114| 1.73| 3.98| 4.14| 4.30| 4.31| 6.52|
|115| 2.27| 2.85| 1.90| 2.97| 1.74| 0.00|
|116| 4.16| 4.17| 5.02| 4.56| 3.45| 3.02|
|117| 4.78| 5.89| 4.60| 5.14| 3.98| 5.74|
|118| 5.52| 6.19| 4.71| 5.02| 3.43| 3.57|
|119| 5.59| 6.86| 4.68| 5.03| 3.59| 3.37|
|120| 4.26| 5.05| 5.04| 4.55| 3.21| 0.00|
|121| 3.99| 4.87| 5.08| 4.56| 3.07| 0.00|
|122| 4.51| 4.49| 5.01| 4.93| 3.75| 2.00|
|123| 5.49| 5.78| 4.13| 4.55| 3.23| 3.55|
|124| 4.18| 4.84| 4.52| 4.50| 3.66| 5.68|
|125| 4.92| 5.60| 4.49| 4.58| 2.77| 1.55|
|126| 4.57| 4.80| 5.04| 4.92| 2.63| 1.53|
|127| 4.52| 4.83| 4.90| 4.81| 2.60| 2.00|
|128| 4.43| 4.56| 4.95| 4.60| 2.43| 1.88|
|129| 3.88| 4.67| 3.30| 3.25| 1.87| 0.83|
|130| 3.71| 4.79| 3.37| 3.14| 1.45| 5.78|
|131| 4.11| 4.84| 4.07| 4.21| 3.63| 4.50|
|132| 4.92| 5.50| 4.04| 4.32| 2.50| 1.04|
|133| 6.33| 7.04| 5.99| 6.50| 3.30| 2.72|
|134| 7.45| 7.25| 6.48| 6.17| 3.97| 2.12|
|135| 6.69| 7.10| 6.36| 6.54| 3.34| 2.21|
|136| 6.62| 7.33| 6.97| 5.91| 3.48| 6.21|
|137| 6.96| 7.47| 5.17| 5.11| 3.35| 3.94|
|138| 6.14| 8.18| 6.18| 6.73| 5.42| 4.19|
|139| 7.18| 9.70| 6.17| 6.86| 4.54| 1.19|
|140| 5.07| 5.92| 4.96| 5.13| 4.08| 4.19|
|141| 4.55| 5.88| 5.38| 4.80| 4.63| 2.23|
|142| 7.81| 9.66| 8.04| 7.14| 4.30| 3.98|
|143*| 8.87| 10.39| 8.22| 7.50| 3.48| 5.61|
|144| 4.70| 5.72| 4.53| 4.93| 4.06| 2.59|
|145| 5.20| 6.29| 4.67| 4.95| 3.38| 4.21|
|146| 6.58| 6.70| 6.10| 5.79| 4.25| 2.21|
|147| 6.35| 6.90| 5.51| 5.95| 4.98| 5.53|
|   |   |   |   |   |   |
|---|---|---|---|---|---|
| 148 | 6.66 | 7.13 | 6.10 | 5.91 | 3.46 |
| 149 | 4.73 | 6.41 | 4.51 | 5.23 | 3.01 |
| 150 | 6.46 | 6.95 | 6.81 | 5.81 | 3.67 |
| 151 | 6.13 | 6.56 | 4.97 | 5.12 | 3.51 |
| 152 | 4.95 | 5.07 | 4.62 | 4.61 | 2.62 |
| 153 | 7.45 | 8.27 | 6.90 | 6.27 | 3.29 |
| 154 | 5.97 | 7.05 | 6.30 | 6.29 | 5.52 |
| 155 | 6.56 | 7.79 | 6.51 | 6.45 | 5.42 |
| 156 | 6.87 | 7.28 | 8.77 | 6.74 | 3.52 |
| 157* | 7.60 | 9.15 | 9.40 | 7.57 | 3.41 |
| 158 | 6.40 | 7.29 | 6.48 | 6.17 | 4.67 |
| 159 | 7.30 | 8.07 | 6.53 | 6.32 | 4.79 |
| 160 | 6.54 | 7.47 | 6.33 | 6.14 | 5.31 |
| 161* | 7.67 | 8.43 | 6.38 | 6.29 | 5.21 |
| 162 | 7.93 | 7.69 | 6.26 | 6.43 | 4.39 |
| 163 | 8.44 | 8.61 | 6.24 | 6.62 | 3.36 |
| 164 | 5.61 | 6.46 | 4.98 | 5.07 | 3.06 |
| 165 | 6.36 | 6.64 | 4.96 | 5.06 | 4.76 |
| 166 | 6.39 | 7.53 | 6.27 | 6.32 | 4.49 |
| 167 | 6.83 | 8.04 | 6.31 | 6.48 | 4.30 |
| 168 | 6.27 | 6.82 | 6.77 | 6.55 | 4.75 |
| 169 | 6.54 | 7.76 | 6.89 | 6.62 | 3.73 |
| 170 | 6.83 | 7.86 | 6.78 | 6.54 | 4.32 |
| 171 | 6.58 | 7.02 | 5.83 | 5.89 | 3.95 |
| 172 | 6.92 | 7.72 | 5.99 | 5.81 | 3.22 |
| 173 | 5.72 | 6.16 | 5.69 | 5.83 | 2.20 |
| 174 | 6.62 | 6.73 | 5.84 | 6.07 | 1.11 |
| 175 | 6.00 | 6.79 | 6.90 | 5.93 | 3.83 |
| 176 | 6.66 | 7.63 | 6.92 | 6.26 | 3.94 |
| 177 | 6.99 | 7.20 | 5.59 | 6.07 | 5.97 |
| 178 | 7.67 | 8.06 | 5.96 | 6.33 | 5.15 |
| 179 | 7.30 | 8.04 | 5.76 | 6.24 | 4.80 |
| 180 | 7.69 | 7.96 | 5.71 | 6.19 | 4.65 |
| 181 | 6.74 | 7.07 | 5.56 | 5.84 | 4.60 |
| 182 | 7.41 | 8.01 | 5.70 | 5.95 | 4.83 |
| 183 | 7.20 | 7.55 | 5.69 | 5.80 | 2.20 |
| 184 | 7.17 | 7.56 | 5.70 | 5.99 | 2.20 |
| 185 | 6.33 | 7.18 | 6.26 | 5.97 | 5.19 |
| 186 | 7.15 | 8.13 | 6.43 | 6.16 | 3.99 |
| 187 | 5.01 | 5.23 | 6.26 | 5.46 | 3.64 |
| 188 | 6.09 | 6.67 | 6.67 | 6.45 | 4.50 |
| 189 | 6.04 | 7.73 | 6.82 | 6.53 | 2.56 |
|   |   |   |   |   |   |
|---|---|---|---|---|---|
| 190 | 7.93 | 8.66 | 6.51 | 6.69 | 5.92 | 4.35 | 6.68 | 6419835 |
| 191* | 8.76 | 8.82 | 6.65 | 6.64 | 5.57 | 7.19 | 7.27 | 6419835 |
| 192 | 8.74 | 9.52 | 6.62 | 6.82 | 4.04 | 3.50 | 6.54 | 6419835 |
| 193 | 7.77 | 7.86 | 7.60 | 7.11 | 3.18 | 3.00 | 6.09 | 6451060 |
| 194 | 5.85 | 7.19 | 4.76 | 4.99 | 6.29 | 3.85 | 5.49 | 9548634 |
| 195 | 5.01 | 7.06 | 5.29 | 4.51 | 6.24 | 4.51 | 5.48 | 9548634 |
| 196 | 7.80 | 8.60 | 6.57 | 6.75 | 4.88 | 2.82 | 6.24 | 10095770 |
| 197 | 5.71 | 7.61 | 6.74 | 6.74 | 3.17 | 2.89 | 5.48 | 10386850 |
| 198 | 6.74 | 8.51 | 6.86 | 6.90 | 4.17 | 4.94 | 6.35 | 10386850 |
| 199 | 7.38 | 8.82 | 6.86 | 6.89 | 4.10 | 2.78 | 6.14 | 10386850 |
| 200 | 5.64 | 5.66 | 6.44 | 5.45 | 3.62 | 6.67 | 5.58 | 10955174 |
| 201 | 5.29 | 6.03 | 5.75 | 5.99 | 3.50 | 1.62 | 4.70 | 11165077 |
| 202 | 5.95 | 5.76 | 5.66 | 5.34 | 4.11 | 2.54 | 4.89 | 11223782 |
| 203 | 4.97 | 5.67 | 5.65 | 5.90 | 3.33 | 5.93 | 5.24 | 11673265 |
| 204 | 5.77 | 5.95 | 6.80 | 5.83 | 4.44 | 4.57 | 5.56 | 11827150 |
| 205 | 6.22 | 7.26 | 6.52 | 6.59 | 5.29 | 4.52 | 6.07 | 11827150 |
| 206 | 5.35 | 8.41 | 6.45 | 5.97 | 3.98 | 2.25 | 5.40 | 11827150 |
| 207 | 4.96 | 6.23 | 5.71 | 5.70 | 4.46 | 3.09 | 5.02 | 16091559 |
| 208 | 5.83 | 6.48 | 6.49 | 6.24 | 5.19 | 4.45 | 5.78 | 44258361 |
| 209 | 5.48 | 6.42 | 6.43 | 6.01 | 5.17 | 4.56 | 5.68 | 44258361 |
| 210 | 5.40 | 6.44 | 6.44 | 6.30 | 5.03 | 3.87 | 5.58 | 44258361 |
| 211 | 5.30 | 6.28 | 6.55 | 6.25 | 4.73 | 4.32 | 5.57 | 44258361 |
| 212 | 7.18 | 8.13 | 8.59 | 7.32 | 5.63 | 5.56 | 7.07 | 44421210 |
| 213 | 7.28 | 7.93 | 8.70 | 6.72 | 5.33 | 4.28 | 6.71 | 44421210 |
| 214 | 8.39 | 8.61 | 8.45 | 7.19 | 4.17 | 3.56 | 6.73 | 44421210 |
| 215 | 7.75 | 8.23 | 8.57 | 6.80 | 3.63 | 3.27 | 6.37 | 44421210 |
| 216 | 7.06 | 8.18 | 8.22 | 7.10 | 3.72 | 2.55 | 6.14 | 52947057 |
| 217 | 6.68 | 8.33 | 8.19 | 6.48 | 3.56 | 4.16 | 6.23 | 52947057 |
| 218 | 5.26 | 5.38 | 5.05 | 4.84 | 3.47 | 6.13 | 5.02 | 62379750 |
| 219 | 5.83 | 5.98 | 5.33 | 4.90 | 3.56 | 5.24 | 5.14 | 62379750 |
| 220 | 4.43 | 5.09 | 5.01 | 5.12 | 3.39 | 4.91 | 4.66 | 92023653 |
| 221 | 6.68 | 7.27 | 5.98 | 5.91 | 4.13 | 4.14 | 5.69 | 101918993 |
| 222 | 7.10 | 7.62 | 7.76 | 7.30 | 3.26 | 2.69 | 5.95 | 4970 |
| 223 | 7.16 | 7.73 | 8.06 | 7.20 | 2.94 | 1.20 | 5.32 | 4970 |
| 224 | 7.81 | 8.52 | 7.86 | 7.26 | 3.10 | 5.51 | 6.68 | 4970 |
| 225 | 3.89 | 4.28 | 3.38 | 3.39 | 2.33 | 1.48 | 3.19 | 11620 |
| 226 | 4.02 | 4.58 | 3.20 | 3.57 | 2.32 | 1.14 | 3.20 | 11620 |
| 227 | 3.90 | 4.32 | 3.15 | 3.46 | 2.30 | 3.35 | 3.05 | 11620 |
| 228 | 4.14 | 4.63 | 3.29 | 3.48 | 3.27 | 2.74 | 3.53 | 11620 |
| 229 | 5.83 | 5.73 | 4.98 | 5.17 | 3.48 | 2.58 | 4.62 | 16871 |
| 230 | 6.41 | 7.56 | 5.85 | 5.68 | 3.46 | 78160 |
| 231 | 3.26 | 3.76 | 2.63 | 2.78 | 1.76 | 6.59 | 5.26 | 10682896 |
|   | 232 | 6.17 | 6.74 |   | 233 | 5.94 | 7.27 |   | 234 | 6.74 | 7.73 |   | 235 | 7.78 | 7.92 |   | 236 | 3.81 | 4.68 |   | 237 | 3.95 | 4.19 |   | 238 | 3.38 | 3.79 |   | 239 | 6.33 | 6.55 |   | 240 | 3.67 | 4.25 |   | 241 | 6.70 | 8.07 |   | 242 | 5.30 | 6.26 |   | 243 | 4.53 | 6.28 |   | 244 | 6.27 | 6.50 |   | 245 | 5.80 | 6.86 |   | 246 | 6.10 | 6.42 |   |
|---|-----|-----|-----|---|-----|-----|-----|---|-----|-----|-----|---|-----|-----|-----|---|-----|-----|---|-----|-----|---|-----|-----|---|-----|-----|---|-----|-----|---|-----|-----|---|-----|-----|---|-----|-----|---|-----|-----|---|-----|-----|---|
|   | 6.83 | 6.73 |   | 4.02 | 1.08 |   | 5.26 | 179806 |   | 6.22 | 6.80 |   | 3.82 | 1.64 |   | 5.45 | 197835 |   | 7.10 | 6.46 |   | 3.12 | -0.26 |   | 5.15 | 197835 |   | 6.78 | 6.73 |   | 3.50 | 1.46 |   | 5.69 | 197835 |   | 3.40 | 3.61 |   | 2.33 | 0.50 |   | 3.05 | 206035 |   | 3.45 | 3.60 |   | 1.91 | 2.92 |   | 3.34 | 206035 |   | 3.07 | 3.10 |   | 2.03 | 2.95 |   | 3.05 | 206035 |   | 6.70 | 6.09 |   | 3.51 | 3.96 |   | 5.52 | 442361 |   | 3.48 | 3.56 |   | 2.33 | 7.78 |   | 4.18 | 3080557 |   | 3.82 | 6.81 |   | 3.21 | 3.80 |   | 6.15 | 5281331 |   | 5.38 | 4.68 |   | 3.11 | 0.00 |   | 4.12 | 5317303 |   | 5.35 | 4.72 |   | 2.87 | 2.99 |   | 4.46 | 5319779 |   | 5.94 | 6.19 |   | 3.65 | 5.31 |   | 5.64 | 14466152 |   | 6.17 | 5.35 |   | 4.50 | 5.17 |   | 5.64 | 11848147 |   | 6.65 | 5.97 |   | 3.97 | 6.06 |   | 5.86 | 11848147 |

**Note:** The virtual screening in Schrödinger was conducted using Glide (10.5) module which invokes ChemScore empirical scoring function.

\[
\Delta G_{\text{binding}} = \Delta G_{\text{match}} F_{\text{match}} + \Delta G_{\text{lipo}} F_{\text{lipo}} + \Delta G_{\text{ambig}} F_{\text{ambig}} + \Delta G_{\text{clash}} F_{\text{clash}} + \Delta G_{\text{rot}} n_{\text{rot}}
\]

\(\Delta G_i\) are coefficients and \(F_i\) are functions

\(\Delta G_{\text{match}}\) is a sum of scores for directed interactions between receptor and ligand.

The terms \(F_{\text{match}}\) provide a measure of hydrophobic contact surface as functions of atomic pairs in receptor and ligands, \(F_{\text{lipo}}\) involving only pairs of unpolar atoms and \(F_{\text{ambig}}\) involving pairs of one polar and one unpolar atom. Finally, \(F_{\text{clash}}\) is a penalty function for protein ligand overlap, and \(n_{\text{rot}}\) is equal to the number of rotatable bonds in the ligand times a weighting factor by default is set as 1. The term \(\Delta G_{\text{rot}} n_{\text{rot}}\) was originally intended as a measure of the entropic cost of freezing intramolecular degrees of freedom in the ligand upon binding with receptor. Virtual screening, mainly serves to suppress the dependence of the score on the molecular weight.\(^{11}\)

The natural compound database was subjected to screening based on extra precision (XP) docking against minimized conformations of Hup receptor created grid.\(^{12}\)

**XP Glide Score =** \(E_{\text{coul}} + E_{\text{vdW}} + E_{\text{bind}} + E_{\text{penalty}}\)

\(E_{\text{bind}} = E_{\text{hyd-enclosure}} + E_{\text{hb-na-motif}} + E_{\text{hb-cc-motif}} + E_{\text{pl}} + E_{\text{hb-pair}} + E_{\text{phobic-pair}}\)

\(E_{\text{penalty}} = E_{\text{desolvation}} + E_{\text{ligand-strain}}\)

\(E_{\text{hyd-enclosure}}\) Represents Hydrophobic Interaction, \(E_{\text{hb}}\) represents hydrogen bonds.
All 246 molecular structures were subjected to Induced-fit standard precision docking in which hydroxyl groups are allowed to reorient thus allowing hydrogen bonding. For each ligand 400 poses we generated and van der Waals scaling of 0.4 Å was allocated on non-polar atoms of both protein and ligand. Keeping prime energy values within 20 kcal/mol, the ligands were re-docked on prime refined structure whereby minimizing the docked pose within the radius of 5 Å.13
**Figure S4:** Long run MD-simulation performed in explicit water solvent for ligand free and bound forms of monomeric Hup (HupM) and dimeric Hup (HupD). **(A)** A representative ray-traced picture of simulated system with simulation cell boundaries set to periodic. Atoms that stick out of the simulation cell will be wrapped to the opposite side of the cell during the simulation. **(B)** The table showing components of the system and time of simulation. **(C, D)** The backbone conformations of ligand free HupD and HupM structures compared before and after the MD simulation. **(E-G)** The docked conformations of Hup-EGCG complex compared before and after the long run MD simulations.
**Figure S5**: The trajectory analysis of MD-simulation data (on free Hup) performed in three blocks. The plots in (A, B and C) and (A', B' and C') show the results, respectively, on HupD and HupM: (A,A') total potential energy of the system, (B,B') number of hydrogen bonds between solute and solvent and (C,C') protein secondary structure content as a function of simulation time.
Figure S6: The trajectory analysis of MD-simulation data (on free Hup) performed in three blocks. The plots in (A, B and C) and (A’, B’ and C’) show the results, respectively, on HupD and HupM: (A,A’) Radius of
gyration of protein in angstrom (Å), (B,B') backbone Root mean square deviation (RMSD) (C,C') Root mean square fluctuations (RMSF) during the simulation plotted as a function of residue number.
Figure S7: (A,B) Time evolution of protein residue secondary structure type during the MD simulation on free Hup (trajectory analysis was performed in three blocks)  (C,D) The time evolution of number of contacts per residue as to see how densely a certain residue range is packed and allows identifying structurally very important residues, e.g. a phenylalanine in the hydrophobic core can contact 15 or more other residues.
Per-residue secondary structure

Helix  Sheet  Turn  Sheet  Helix3_{10}  HelixPi

Residue Number

Simulation time (in nanosecond)

Per-residue number of contacts
Figure S8: The results obtained from the trajectory analysis of MD simulation data on top two best complexes of EGCG with HupM. (A, A') Total potential energy of the system plotted as a function of simulation time. (B, B') The plot shows the number of hydrogen bonds between solute and solvent as a function of simulation time. (C, C') The plot shows time evolution of protein secondary structure content.
Figure S9: The results obtained from the trajectory analysis of MD simulation data on top two best complexes of EGCG with HupM. (A,A’) Radius of gyration plotted as a function of simulation time. (B, B’) Protein RMSD from the starting structure plotted as a function of simulation time. (C,C’) The protein residue-wise root mean square fluctuation (RMSF) calculated from the average RMSF of the atoms constituting the residue.
Figure S10: The results obtained from the trajectory analysis of MD simulation data on top two best complexes of EGCG with HupM. (A,A') The ligand movement RMSD measured after superposing the receptor on its reference structure. This procedure delivers information about the movement of the ligand in its binding pocket. (B, B') The ligand conformation RMSD after superposing on the ligand plotted as a function of simulation time. This provides information about the conformational changes of the ligand. (C,C') The time evolution of number of contacts per residue as to see how densely a certain residue range is packed and allows identifying structurally very important residues, e.g. a phenylalanine in the hydrophobic core can contact 15 or more other residues.
Figure S11: The results obtained from the trajectory analysis of long run MD simulation data on EGCG-HupD complex with highest binding energy. **(A)** Total potential energy of the system plotted as a function of simulation time. **(B)** The plot shows the number of hydrogen bonds between solute and solvent as a function of simulation time. **(C)** The plot shows protein secondary structure content as a function of simulation time. **(D)** The ligand movement RMSD measured after superposing the receptor on its reference structure. This procedure delivers information about the movement of the ligand in its binding pocket. **(E)** The ligand conformation RMSD after superposing on the ligand plotted as a function of simulation time. This provides information about the conformational changes of the ligand.
Figure S12: The results obtained from the trajectory analysis of MD simulation data on ECGC-HupD complex with highest binding energy. (A) Radius of gyration as a function of simulation time. (B) Protein RMSD from the starting structure plotted as a function of simulation time. (C,C') The plot shows time evolution of number of contacts per residue as to see how densely a certain residue range is packed and allows identifying structurally very important residues. (D) The protein residue-wise root mean square fluctuation (RMSF) calculated from the average RMSF of the atoms constituting the residue.
Figure S13: Time evolution of protein residue secondary structure type during the MD simulation (trajectory analysis was performed in three blocks): (A) Chain-A of HupD in complex with EGCG, (B) Chain-B of HupD in complex with EGCG, (C) Highest energy complex of EGCG with HupM and (D) second highest energy complex of EGCG with HupM.
**Appendix-II: **$^1$H and $^{13}$C NMR Assignment of Epigallocatechin-Gallate (EGCG):

All experiments were performed on a Bruker Avance (III) 800 MHz NMR spectrometer equipped with a 5 mm inverse detection cold probe and an actively shielded z-gradient and operating at $^1$H frequency of 800 MHz and $^{13}$C frequency of 200 MHz. All spectra were measured at temperature of 300 K using 100% D$_2$O as solvent. The conventional one dimensional (1D) $^1$H (at 800 MHz frequency) and $^{13}$C (at 200 MHz frequency) NMR spectra were recorded using the standard pulse programs from Bruker library. The various acquisition parameters and the names of the Bruker pulse programs are summarized in **Table S2**.

**Table S2:** The acquisitions and processing parameters used for NMR experiments performed to confirm the resonance assignment of Epigallocatechin-Gallate (EGCG).

| Parameters          | RF pulse Offsets (ppm) | Spectral widths (ppm) | Size of FID (direct/indirect) | Zero-filling before FT | Recycle delay | Number of scans per FID |
|---------------------|------------------------|-----------------------|-------------------------------|------------------------|---------------|-------------------------|
| 1D $^1$H NMR (zg)   | 4.7                    | 20                    | 64 k                          | 64 k                   | 1.0 sec       | 16                      |
| 1D $^{13}$C NMR (zgpg30) | 80                    | 265                   | 64 k                          | 64 k                   | 1.0 sec       | 2048                    |

Note: RF: Radiofrequency; 1k=1024 points; 1D: one dimensional; FID: Free Induction decay
Figure S14: The 1D (A) $^1$H NMR and (B) $^{13}$C NMR spectra of purified Epigallocatechin-Gallate (EGCG, dissolved in 100% D$_2$O) recorded at 800 MHz NMR spectrometer.

Table S3: The NMR samples prepared for studying the interaction between Epigallocatechin-Gallate (EGCG, MW=458.37 g/mol) and Hup. The concentrations of prepared stock solutions of Hup are 600 μM (i.e. 0.6 mM) and of EGCG are 20 mM (in 100% D$_2$O) respectively. The final concentration of EGCG in each sample was 100 μM (also checked using UV spectrometry based on its extinction coefficient (9700 M$^{-1}$ cm$^{-1}$ at 280 nm)).

| #  | NMR sample Final Volume (μL) | 0.6 mM Hup Solution in μL | EGCG solution of 20mM | Sodium phosphate Buffer solution | Final Hup Conc. |
|----|-----------------------------|---------------------------|----------------------|----------------------------------|-----------------|
| 1  | 600                         | 0                         | 3 μl                 | 597                              | 0.0 μM          |
| 2  | 600                         | 4                         | 3 μl                 | 593                              | 4.0 μM          |
| 3  | 600                         | 8                         | 3 μl                 | 589                              | 8.0 μM          |
| 4  | 600                         | 12                        | 3 μl                 | 585                              | 12.0 μM         |
| 5  | 600                         | 16                        | 3 μl                 | 581                              | 16.0 μM         |
| 6  | 600                         | 20                        | 3 μl                 | 577                              | 20.0 μM         |
Table S4A: The signal to noise ratio analysis for different NMR peaks of Epigallocatechin-Gallate (EGCG) sample (Concentration~ 100 μM) recorded in the absence and presence of Hup protein (0.0, 4.0, 8.0, 12.0, 16.0 and 20.0 μM). \( I_0 \) = Ligand EGCG signal in its free form; \( I \) = Ligand signal after the addition of protein; \( I/I_0 \) = normalized value; \( \theta = 1 - (I/I_0) \) is the degree of binding; TPA: Three point average.

| Hup Concentration in micromolar | Signal Intensity \((I)\) | Initial Signal \((I_0)\) | NS \(I/I_0\) | \( \theta \) (degree of binding) | Three point averaging |
|---------------------------------|-------------------------|------------------------|--------------|-------------------------------|----------------------|
| **EGCG Signal at 7.09 ppm**     |                         |                        |              |                               |                      |
| 0                               | 474.92                  | 474.92                 | 1.00         | 0.00                          | 0.00                 |
| 4                               | 209.34                  | 474.92                 | 0.44         | 0.56                          | 0.44                 |
| 8                               | 118.34                  | 474.92                 | 0.25         | 0.75                          | 0.71                 |
| 12                              | 92.05                   | 474.92                 | 0.19         | 0.81                          | 0.80                 |
| 16                              | 81.56                   | 474.92                 | 0.17         | 0.83                          | 0.85                 |
| 20                              | 45.31                   | 474.92                 | 0.10         | 0.90                          | 0.87                 |
| **EGCG Signal at 6.70 ppm**     |                         |                        |              |                               |                      |
| 0                               | 1178.98                 | 1178.98                | 1.00         | 0.00                          | 0.00                 |
| 4                               | 168.11                  | 1178.98                | 0.14         | 0.86                          | 0.58                 |
| 8                               | 147.29                  | 1178.98                | 0.12         | 0.88                          | 0.88                 |
| 12                              | 96.19                   | 1178.98                | 0.08         | 0.92                          | 0.91                 |
| 16                              | 70.10                   | 1178.98                | 0.06         | 0.94                          | 0.94                 |
| 20                              | 37.29                   | 1178.98                | 0.03         | 0.97                          | 0.95                 |
| **EGCG Signal at 6.28 ppm**     |                         |                        |              |                               |                      |
| 0                               | 117.77                  | 117.77                 | 1.00         | 0.00                          | 0.00                 |
| 4                               | 98.34                   | 117.77                 | 0.84         | 0.16                          | 0.21                 |
| 8                               | 64.52                   | 117.77                 | 0.55         | 0.45                          | 0.39                 |
| 12                              | 53.78                   | 117.77                 | 0.46         | 0.54                          | 0.59                 |
| 16                              | 26.50                   | 117.77                 | 0.23         | 0.77                          | 0.73                 |
| 20                              | 14.23                   | 117.77                 | 0.12         | 0.88                          | 0.83                 |
| **EGCG Signal at 6.25 ppm**     |                         |                        |              |                               |                      |
| 0                               | 110.00                  | 110.00                 | 1.00         | 0.00                          | 0.00                 |
| 4                               | 106.86                  | 110.00                 | 0.97         | 0.03                          | 0.12                 |
| 8                               | 72.55                   | 110.00                 | 0.66         | 0.34                          | 0.27                 |
| 12                              | 59.91                   | 110.00                 | 0.54         | 0.46                          | 0.51                 |
| 16                              | 30.67                   | 110.00                 | 0.28         | 0.72                          | 0.67                 |
| 20                              | 16.82                   | 110.00                 | 0.15         | 0.85                          | 0.78                 |
| **EGCG Signal at 5.73 ppm**     |                         |                        |              |                               |                      |
| 0                               | 159.05                  | 159.05                 | 1.00         | 0.00                          | 0.00                 |
| 4                               | 75.58                   | 159.05                 | 0.48         | 0.52                          | 0.12                 |
| 8                               | 44.29                   | 159.05                 | 0.28         | 0.72                          | 0.27                 |
| 12                              | 46.07                   | 159.05                 | 0.29         | 0.71                          | 0.51                 |
| 16                              | 25.80                   | 159.05                 | 0.16         | 0.84                          | 0.67                 |
| 20                              | 20.61                   | 159.05                 | 0.13         | 0.87                          | 0.78                 |
### Supplementary Material

|       | EGCG Signal at 5.27 ppm |       | EGCG Signal at 3.18 ppm |       | EGCG Signal at 3.04 ppm |
|-------|-------------------------|-------|-------------------------|-------|-------------------------|
| 0     | 247.04                  | 1.00  | 181.10                  | 1.00  | 223.61                  |
| 4     | 73.43                   | 0.30  | 38.36                   | 0.21  | 50.61                   |
| 8     | 47.41                   | 0.19  | 23.30                   | 0.13  | 32.38                   |
| 12    | 44.82                   | 0.18  | 17.08                   | 0.09  | 23.05                   |
| 16    | 22.52                   | 0.09  | 11.99                   | 0.07  | 16.34                   |
| 20    | 10.19                   | 0.04  | 11.33                   | 0.06  | 14.81                   |

### Table: EGCG Signal at 5.27 ppm

- Time (min): 0
- Signal (ppm): 247.04
- Amplitude: 1.00
- Area: 0.00
- Peak: 0.00

### Table: EGCG Signal at 3.18 ppm

- Time (min): 0
- Signal (ppm): 181.10
- Amplitude: 1.00
- Area: 0.00
- Peak: 0.00

### Table: EGCG Signal at 3.04 ppm

- Time (min): 0
- Signal (ppm): 223.61
- Amplitude: 1.00
- Area: 0.00
- Peak: 0.00
Table S4B: The average value of degree of binding ($\theta$) estimated from the individual three point averaged values of $\theta$ obtained for different NMR signals of Epigallocatechin-Gallate (EGCG) for evaluating the apparent dissociation constant (kD) for the interaction between EGCG and Hup.

| Hup Conc. in micromolar | The degree of binding ($\theta$) for individual signals of EGCG (in ppm) and its average value |
|-------------------------|-------------------------------------------------------------------------------------------------|
|                         | 7.09 | 6.70 | 6.28 | 6.25 | 5.73 | 5.27 | 3.18 | 3.04 | Average |
| 0                       | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00     |
| 4                       | 0.44 | 0.58 | 0.21 | 0.12 | 0.12 | 0.50 | 0.55 | 0.54 | 0.46     |
| 8                       | 0.71 | 0.88 | 0.39 | 0.27 | 0.27 | 0.78 | 0.86 | 0.84 | 0.73     |
| 12                      | 0.80 | 0.91 | 0.59 | 0.51 | 0.51 | 0.85 | 0.90 | 0.89 | 0.81     |
| 16                      | 0.85 | 0.94 | 0.73 | 0.67 | 0.67 | 0.90 | 0.93 | 0.92 | 0.87     |
| 20                      | 0.87 | 0.95 | 0.83 | 0.78 | 0.78 | 0.93 | 0.94 | 0.93 | 0.91     |

Note: The average values do not include the NMR signal intensities for EGCG peaks at 6.28 and 6.25 ppm because their intensity profiles were found deviating from the hyperbolic model upon saturating with protein Hup.
Appendix-III:

Determination of the binding constant from saturation binding curve based on NMR titration experiments

For the single binding site protein ligand interaction system (i.e. for a 1:1 complex), the equilibrium involved is described according to:\(^\text{(S3)}\)

\[
P + L \underset{k_{\text{off}}}{\overset{k_{\text{on}}}{\rightleftharpoons}} PL
\]

\[k_D = \frac{[P][L]}{[PL]} = \frac{k_{\text{off}}}{k_{\text{on}}} \quad \text{(S4)}\]

where [P], [L], and [PL] are equilibration concentrations of the receptor, ligand, and the receptor-ligand complex, respectively. The size of dissociation constant (\(K_D\)) is determined by the on (\(k_{\text{on}}\)) and off (\(k_{\text{off}}\)) rates of the ligand on its target and has the units of concentration.\(^\text{(S5)}\)

Theoretically, a value of \(K_D\) in the mM range implies an approximately 1:1000 ratio of free to bound states of ligand in an equimolar mixture of P and L and a KD in the \(\mu\)M range implies an approximately 1:10,000 ratio of these states, i.e., a much more stable complex with less of the ‘free’ species present.

Now, the dissociation constant can be written by introducing total ligand concentration, \(L_{\text{tot}}\), as follows:

\[k_D = \frac{[P][L]}{[PL]} \quad \text{(S5)}\]

\[k_D[PL] = [P][L] = [P](L_{\text{tot}} - [PL])\]

\[k_D[PL] = [P]L_{\text{tot}} - [PL][P]\]

\[k_D[PL] + [PL][P] = [P]L_{\text{tot}}\]

\[\frac{[PL]}{L_{\text{tot}}} = \frac{k_D}{k_D + [P]}\]

\[
L_{\text{tot}} = \frac{k_D + [P]}{k_D} = \frac{[P]}{k_D + [P]}
\]

The equation describes a hyperbola and for a ligand based NMR titration, this can be described as

\[
\frac{[P]}{k_D + [P]} = \frac{[PL]}{L_{\text{tot}}} = \frac{I_0 - I}{I_0} = \theta \quad \text{(S6)}
\]

Where \(L_{\text{tot}}\) is the concentration of total free ligand evaluated based on \(I_0\) (i.e. initial signal intensity of ligand in its free form) and \(I\) is the signal intensity of ligand during the titration in the absence and presence of Hup. The derivation of above equation is based on the assumption that there is no signal contribution
from bound ligand or there is no signal attenuation of ligand due to its interaction with receptor. As both these assumptions are very crude, therefore, the estimated dissociation constants has been referred here as apparent dissociation constants.
Appendix-IV: Virtual screening of higher molecular weight (HMW) library of AHP natural compound against the protein receptor Hup

Table S5: The binding energy parameters derived from virtual screening of HMW compound library against dimeric and monomeric Hup structures employing small molecule docking approaches (AUTODOCK, VINA and Glide).

| HMW Compound | AUTODOCK | VINA | Glide | Cumulative Average BE |
|--------------|----------|------|-------|-----------------------|
| LID | CID | HupM | HupD | HupM | HupD | HupM | HupD | HupM | HupD |
| 1 | 9852086 | 5.25 | 5.25 | 7.66 | 6.63 | 6.40 | 3.64 | 5.81 |
| 2 | 73178 | 8.03 | 8.03 | 7.07 | 7.69 | 6.12 | 10.73 | 7.95 |
| 3 | 73178 | 9.20 | 9.20 | 7.23 | 7.65 | 7.82 | 10.98 | 8.68 |
| 4 | 73568 | 8.79 | 8.79 | 7.05 | 6.99 | 6.93 | 11.70 | 8.38 |
| 5 | 73568 | 8.71 | 8.71 | 7.04 | 7.21 | 5.85 | 10.65 | 8.03 |
| 6 | 114627 | 8.08 | 8.08 | 6.70 | 7.43 | 7.22 | 7.25 | 7.46 |
| 7 | 442428 | 7.63 | 7.63 | 7.06 | 7.01 | 7.24 | 6.41 | 7.16 |
| 8 | 442439 | 7.09 | 7.09 | 6.89 | 7.28 | 7.37 | 7.73 | 7.24 |
| 9 | 5280805 | 8.41 | 8.41 | 6.92 | 7.31 | 7.19 | 9.58 | 7.97 |
| 10 | 5280805 | 9.11 | 9.11 | 6.97 | 7.33 | 6.78 | 6.48 | 7.63 |
| 11 | 5281800 | 7.34 | 7.34 | 6.94 | 6.68 | 7.16 | 6.46 | 6.99 |
| 12 | 5281847 | 8.83 | 8.83 | 8.57 | 7.09 | 3.85 | 8.90 | 7.68 |
| 13 | 5281847 | 9.63 | 9.63 | 8.47 | 7.38 | 3.76 | 5.64 | 7.42 |
| 14 | 5282153 | 11.04 | 11.04 | 7.00 | 7.31 | 7.28 | 9.84 | 8.92 |
| 15 | 5388496 | 11.72 | 11.72 | 7.30 | 8.08 | 0.00 | 5.21 | 7.34 |
| 16 | 5388496 | 9.94 | 9.94 | 7.84 | 7.85 | 0.00 | 0.00 | 5.93 |
| 17 | 6439941 | 9.03 | 9.03 | 7.72 | 6.70 | 6.74 | 4.60 | 7.30 |
| 18 | 6476333 | 8.06 | 8.06 | 7.21 | 7.80 | 7.78 | 9.27 | 8.03 |
| 19 | 10033935 | 10.80 | 10.80 | 7.13 | 7.98 | 0.00 | 12.80 | 8.25 |
| 20 | 10033935 | 10.33 | 10.33 | 7.01 | 8.09 | 0.00 | 9.36 | 7.52 |
| 21 | 16129778 | 2.56 | 2.56 | 4.32 | 5.31 | 0.00 | 15.64 | 3.46 |
| 22 | 16129778 | 0.79 | 0.79 | 1.56 | 1.43 | 0.00 | 13.56 | 2.52 |
| 23 | 24847856 | 4.96 | 4.96 | 4.98 | 6.08 | 6.13 | 6.07 | 5.53 |
| 24 | 24847856 | 4.61 | 4.61 | 5.06 | 5.30 | 4.87 | 5.93 | 5.06 |
| 25 | 44584733 | 9.74 | 9.74 | 8.07 | 8.12 | 0.00 | 0.00 | 5.94 |
| 26 | 44584733 | 9.78 | 9.78 | 8.04 | 7.95 | 0.00 | 0.00 | 5.92 |
| 27 | 71436711 | 5.68 | 5.68 | 6.69 | 5.05 | 5.37 | 6.34 | 5.80 |
| 28 | 71436711 | 5.50 | 5.50 | 6.65 | 4.96 | 3.87 | 6.28 | 5.46 |
| 29 | 101304443 | 5.92 | 5.92 | 6.82 | 5.14 | 5.89 | 7.14 | 6.14 |
| 30 | 101304443 | 5.36 | 5.36 | 6.77 | 5.10 | 4.72 | 5.58 | 5.48 |
| 31 | 101973939 | 7.07 | 7.07 | 7.63 | 6.61 | 6.55 | 8.23 | 7.19 |
| 32 | 101973939 | 8.26 | 8.26 | 7.78 | 6.76 | 5.35 | 6.37 | 7.13 |
| 33 | 162221834 | 11.46 | 11.46 | 6.44 | 7.42 | 5.54 | 14.37 | 9.45 |
|    | LID     | BE  | BE  | BE  | BE  | BE  | BE  |
|----|---------|-----|-----|-----|-----|-----|-----|
| 34 | 162221834 | 12.27 | 12.27 | 6.56 | 7.59 | 0.60 | 15.20 |
| 35 | 448438  | 7.24 | 7.24 | 8.15 | 7.03 | 3.62 | 0.13 |
| 36 | 5281247 | 7.53 | 7.53 | 8.49 | 7.00 | 3.27 | 1.05 |
| 37 | 12112747 | 7.82 | 7.82 | 9.00 | 7.39 | 5.60 | 2.18 |
| 38 | 12112747 | 8.77 | 8.77 | 8.62 | 7.22 | 4.69 | 0.64 |
| 39 | 23634523 | 6.55 | 6.55 | 7.30 | 5.63 | 5.53 | 4.39 |
| 40 | 23634523 | 7.39 | 7.39 | 7.24 | 5.87 | 5.44 | 4.03 |
| 41 | 23634528 | 7.91 | 7.91 | 7.43 | 6.14 | 5.26 | 4.72 |
| 42 | 23634528 | 6.85 | 6.85 | 7.70 | 5.67 | 5.04 | 4.39 |

Note: LID: Internal ID of compound in the higher molecular weight library; BE: Binding energy (kcal/mol)
Figure S15: (A) Shows the top ten binding hits identified through AUTODOCK (top), VINA (middle) and Schrodinger glide docking using extra-precision algorithm (bottom). The different colours are used to highlight specific compounds present in top hit index. (B) Stacked binding energy (in kcal/mol) obtained after computational screening of HMW library of natural compounds against target structures of HupM and HupD receptors. The top three hit compounds i.e. Gallagic acid (LID: 33,34), Luteolin-7-O-beta-D-diglucuronide (LID: 14) and 1,2,3,6-tetra-O-galloyl-b-D-glucose (LID: 3) are highlighted in (B) and these all three were also found to be present in more than four hit indices shown in (A). Abbreviations used: AHP: Anti-H pylori; HupM: Monomeric Hup; HupD: Dimeric Hup; AD_BE: Autodock binding energy; VINA_BE: VINA binding energy; GS: glide score, LID: Internal ID of compound in Table S15.
Figure S16: (A,B) The best Glide docking poses of Gallagic-acid with HupD and HupM selected after cluster analysis based on highest binding energy of the complex. (C,D) The 2D representation of molecular interaction for Gallagic acid surrounded by contacting receptor residues.
References:

1. Schrödinger, L. version 2.1. *New York, NY* 2009.

2. Raj, R.; Agarwal, N.; Raghavan, S.; Chakraborti, T.; Poluri, K. M.; Kumar, D. Exquisite binding interaction of 18β-Glycyrrhetinic acid with histone like DNA binding protein of Helicobacter pylori: A computational and experimental study. *Int. J. Biol. Macromol.* 2020.

3. Chen, C. C.; Hwang, J. K.; Yang, J. M. (PS)2: Protein structure prediction server. *Nucleic Acids Res.* 2006, 34 (suppl 2), W152-W157.

4. Chen, C. C.; Hwang, J. K.; Yang, J. M. (PS)2-v2: template-based protein structure prediction server. *BMC Bioinf.* 2009, 10 (1), 366.

5. Krieger, E.; Dunbrack, R. L.; Hooft, R. W.; Krieger, B. Assignment of protonation states in proteins and ligands: Combining pK a prediction with hydrogen bonding network optimization. In *Computational Drug Discovery and Design*, Springer: 2012; pp 405-421.

6. Dickson, C. J.; Madej, B. D.; Skjevik, +. A.; Betz, R. M.; Teigen, K.; Gould, I. R.; Walker, R. C. Lipid14: the amber lipid force field. *J. Chem. Theory Comput.* 2014, 10 (2), 865-879.

7. Hornak, V.; Abel, R.; Okur, A.; Strockbine, B.; Roitberg, A.; Simmerling, C. Comparison of multiple Amber force fields and development of improved protein backbone parameters. *Proteins: Struct., Funct., Bioinf.* 2006, 65 (3), 712-725.

8. Krieger, E.; Nielsen, J. E.; Spronk, C. A.; Vriend, G. Fast empirical pKa prediction by Ewald summation. *J. Mol. Graphics Modell.* 2006, 25 (4), 481-486.

9. Krieger, E.; Vriend, G. New ways to boost molecular dynamics simulations. *J. Comput. Chem.* 2015, 36 (13), 996-1007.

10. Chen, D. E.; Willick, D. L.; Ruckel, J. B.; Floriano, W. B. Principal component analysis of binding energies for single-point mutants of hT2R16 bound to an agonist correlate with experimental mutant cell response. *J. Comput. Biol.* 2015, 22 (1), 37-53.

11. Halgren, T. A.; Murphy, R. B.; Friesner, R. A.; Beard, H. S.; Frye, L. L.; Pollard, W. T.; Banks, J. L. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *J. Med. Chem.* 2004, 47 (7), 1750-1759.

12. Friesner, R. A.; Murphy, R. B.; Repasky, M. P.; Frye, L. L.; Greenwood, J. R.; Halgren, T. A.; Sanschagrin, P. C.; Mainz, D. T. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein ligand complexes. *J. Med. Chem.* 2006, 49 (21), 6177-6196.

13. Farid, R.; Day, T.; Friesner, R. A.; Pearlstein, R. A. New insights about HERG blockade obtained from protein modeling, potential energy mapping, and docking studies. *Bioorg. Med. Chem.* 2006, 14 (9), 3160-3173.
14. Fielding, L. NMR methods for the determination of protein-ligand dissociation constants. *Prog. Nucl. Magn. Reson. Spectrosc.* 2007, 51 (4), 219-242.

15. Becker, W.; Bhattiprolu, K. C.; Gubensák, N.; Zangger, K. Investigating Protein-Ligand Interactions by Solution Nuclear Magnetic Resonance Spectroscopy. *ChemPhysChem* 2018, 19 (8), 895-906.