**In-silico** investigation of curcumin drug-likeness, gene-targets and prognostic relevance of the targets in panels of human cancer cohorts

David B Oshevire 1, Aishatu Mustapha 2, Blessing U. Alozieuwa 3, Hassana H. Badeggi 2, Abdulfatai Ismail 4, Opeyemi N. Hassan 5, Peter I Ugwunnaji 6, Jonathan Ibrahim 7, Bashir Lawal 8, * and Eustace B. Berinyu 9

1 Senior Medical Officer, Hayok Medicare, Abuja, Nigeria.
2 Biological Science Department, Niger State Polytechnic Zungeru, Nigeria.
3 Veritas University Abuja, Bwari, FCT-Abuja, P.M.B. 7084.
4 Department of Animal Biology, Federal University of Technology Minna, Nigeria.
5 Centre International Universitaire Des Meilleurs (C.I.U.M), Bestower International University Seme-Podji-Republique Du Benin.
6 Department of Chemistry, Michael Okpara University of Agriculture, Umudike, P.M.B 7267, Umuahia, Abia State, Nigeria
7 Gombe State College of Health Science and Technology, Kaltungo, Nigeria
8 Graduate Institute for Cancer Biology and Drug Discovery, College of Medical Science and Technology, Taipei Medical University, Taipei 11031, Taiwan
9 Faculty of Medicine and Biomedical Science, University of Yaoundé 1, Cameroon.

GSC Biological and Pharmaceutical Sciences, 2021, 14(01), 037–046

Publication history: Received on 05 December 2020; revised on 06 January 2021; accepted on 07 January 2021


**Abstract**

Despite advancements in diagnostic and standard treatment modalities, cancer survival rate remains disappointing globally. It has however, been recognized that exploring the therapeutic properties of secondary metabolite from natural products may alleviate the problems of drug resistance and toxicity that besiege the conventional therapies, and hence improve the overall prognosis of cancer patient. To this end curcumin, a polyphenolic natural compound has been widely studied for its anticancer activities in *in vitro* and *in vivo* models. Computational technology has significantly improved the success rate of drug discovery and development, hence, it has become a widely explore tool in drug candidate identification. In this study we used computational approached to identify 12 genes that are potential druggable candidate for curcumin. The genes identified were found to be enriched in cancer and drug resistance associated signaling pathways. Interestingly, the top 3 identified genes; Microtubule-associated protein tau (MAPT), Toll-like receptor 9 (TLR9) and Tyrosyl-DNA phosphodiesterase 1 (TDP1) were observed to be over expressed in multiple cancer cohorts and were associated with poor prognoses of the patients. Curcumin has good physicochemical, bioavailability and ADMET properties. Importantly, it met the Lipinski’s Rule of 5 for drug likeness and thus worthy of further *in vitro* and *in vivo* confirmation studies.

**Keywords:** Curcumin, Cancer; In silico studies; Drug-likeness; Drug target; Protein-protein interactions.

1. **Introduction**

Despite advancements in diagnostic and standard treatment modalities, cancer remain second leading cause of death worldwide [1]. Currently, there are over 18 million cancer cases with projection of 25 million cases by year 2030 [2,3]. The available treatment modalities for cancers including immunotherapy, radiation therapy and chemotherapy are disappointing, providing a minimal extension of the overall survival and, and general poor prognosis [1]. Furthermore,
the available chemotherapy suffers from drug resistance and serious side effects. The dilemma is further compounded by activations of multiple pathways associated with cancer growth and metastasis, hence necessitating the need to explore natural products for alternative multi-target chemotherapeutic agent with minimal side effect [4].

Natural products are diverse in terms of structure and biological activities [5-8] and thus offer superior therapeutic options and safety than the conventional and synthetic compounds used in high throughput screening processes in drug discovery and development pipelines [9]. Thus, natural products, especially phytochemical extracts from medicinal flora have proven to be immensely valuable as a lead compound for drug discovery and development [10,11]. In fact, an approximated half of the conventional drugs are of phytochemical backbone [12].

Curcumin, is a polyphenolic natural compound isolated from Curcuma longa, a reputable medicinal plants commonly known as turmeric [13]. Curcumin has been reported for several therapeutic activities against a number of diseases; including antioxidants, hepato-protective [14], anti-inflammatory and anticancer activities. Several literatures have emphasized the anti-antineoplastic properties of curcumin against experimental cell line and chemically induced carcinogenesis in animal models of cancer [15-18].

Curcumin have also been reported to synergize and enhance the activities of conventional drugs in combination therapy; studies have revealed that curcumin can sensitize a variety of tumors including glioma, colon cancer, neuroblastoma, epidermal carcinoma, cervical carcinoma, and prostate cancer to chemo and radio therapies [19]. However, dysregulation of several signaling pathways have been implicated in the activities of curcumin [20]. Computational technology has significantly minimized the experimental drug trials and improve the success rate of drug discovery and development, hence, it has become a widely explore tool in drug candidate identification [21]. To this end, we conducted an in silico drug likeness and target predictions studies on curcumin with a view of unraveling the potential targets that could be implicated in it diverse anti-cancer activities.

2. Material and methods

2.1. Medicinal Chemistry, Drug likeness and in-silico ADMET Properties of curcumin

The medicinal chemistry, drug likeness and admet properties of curcumin were evaluated using SwissADME algorithm (http://www.swissadme.ch). The drug likeness of the compounds were predicted in accordance to the Lipinski’s Rule of 5 for drug candidate. The Rule of 5 accommodate a drug candidate with < 5 H-bond donors, < 10 H-bond acceptors, < 500 Da, and < 5 Log P (CLog P) as a good drug candidate [22].

2.2. Identification of curcumin targets, protein-protein interaction network and prognostic relevance of the target genes

The curcumin target genes were identified using SwissTarget algorithm. We used STRING servers (https://string-db.org/) to predict protein-protein interactions (PPIs) and functional protein partners of the identified curcumin targets. We used the Tumor Immune Estimation Resource (TIMER2.0) algorithm (http://timer.cistrome.org/) for a differential expression analysis of MAPT, TLR9 and TDP1 in tumor vs. normal tissues from various cancers in The Cancer Genome Atlas (TCGA) database [23]. We used Kaplan-Meier (K-M) survival plots to analyze overall survival (OS) and disease-free survival (DFS) ratios between the cancer cohorts with high and low MAPT, TLR9 and TDP1 expression profiles from the 9736 tumor samples across 33 cancer types in the Gene Expression Profiling Interactive Analysis (GEPIA) database (http://geopia.cancer-pku.cn/) [24].

2.3. Data analysis

The statistical significance of differentially expressed genes was evaluated using the Wilcoxon test. * p<0.05; ** p<0.01; *** p<0.001. The Kaplan-Meier curve was employ to present the patients’ survival from different cancer cohorts.

3. Results and discussion

3.1. Curcumin has acceptable physicochemical properties and met the requirement for drug likeness.

In-silico approach was recently found to aid identifying a safe and possible drug candidate and also to address the possible molecular mechanism of action of modest molecules with a big protein [21]. It is on this note that we conducted medicinal chemistry, pharmacokinetic (ADMET) studies and “drug-likeness” using Lipinski’s rule of five on curcumin (1E,6E)-1,7-bis(4-hydroxy-3-methoxyphenyl) hepta-1,6-diene-3,5-dione (Figure 1A) to determine their activity within
the human body. Our findings revealed that curcumin has acceptable physicochemical and bioavailability properties (Figure 1C and Table 1). It is passively absorbed by the gastrointestinal tract but has poor blood brain barrier permeation ability (Figure 1B). The ‘Bioavailability Radar’ (Fig. 1c) is designed based on prediction of six physicochemical properties; polarity (POLAR), molecular mass (SIZE), Lipophilicity (LIPO), flexibility (FLEX), saturation (INSATU) and water solubility (INSOLU). These properties of a molecule should fall within desired ranges for it to be suitable as a drug lead. These desired ranges are graphically represented by the pink shaded area of the bioavailability radar. Since all six properties for curcumin are within the pink range, according to the SwissADME calculations, it is suitable to be used as a drug. Also, it was predicted to have a high gastrointestinal absorbance and no violations of rules determining drug-likeness. It has good water solubility (Logs=-3.94) and synthetic availability (2.97%). The results of ADME and drug-likeness properties together with the bioavailability scores suggest that curcumin have good drug-likeness properties, and thus have a good potential to be developed into oral drug for therapeutic application [25].

Figure 1 Curcumin, a multi-target natural product has acceptable physicochemical properties and met the requirement for drug likeness. (A) ring form of 3 dimensional (upper panel) and the 2 dimensional (lower pane) chemical structure of curcumin (B) BOILED-Egg model of brain or intestinal estimated permeation of curcumin. Points located in the BOILED egg yolk are molecule predicted to passively permit via the BBB. Points located in the BOILED egg yolk are molecule predicted to passively absorbed by the gastrointestinal tract. In addition the points in the BOILED egg are coloured-blue if predicted as actively effluxed by P-gp (PGP+) and red-coloured if predicted as non-substrate of P-gp (PGP−). (C); The Bioavailability Radar showing the suitable physicochemical space of oral bioavailability of curcumin. The pink area represents the optimal range of curcumin for each property. LIPO; Lipophilicity, POLAR: Polarity, INSOLU: Insolubility, INSATU: Unsaturation, FLEX: Flexibility. Fraction Csp3; fraction of carbons in the sp3 hybridization. (C) chat showing the classes of gene that are potential targets for curcumin.

3.2. Curcumin are potential multi-targets agent

The past few years have witnessed significant advances in drug discovery technologies for the identification of novel lead compounds against a wide range of therapeutic targets including chemical matter to disrupt protein-protein interactions [26]. In the present study, our drug target analysis of curcumin identified membrane receptors (7%), kinases (20%), enzymes (33%) and others (40%) as the classes of the target genes (figure 1D). Specifically, 12 genes including; Microtubule-associated protein tau, Toll-like receptor 9, Tyrosyl-DNA phosphodiesterase 1, Muscle blind-like protein, Lactoylglutathione lyase, RAC-alpha serine/threonine-protein kinase, Testosterone 17-beta-dehydrogenase 3,
Gamma-secretase C-terminal fragment 59, Quinone oxidoreductase, Estradiol 17-beta-dehydrogenase 12 were identified as a druggable candidate for curcumin with probability range between 0.44–1 (Table 2).

Table 1 Drug likeness, medicinal chemistry and physicochemical and ADMET properties of curcumin

<table>
<thead>
<tr>
<th>Physicochemical Properties</th>
<th>Curcumin value</th>
<th>Reference Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formula</td>
<td>C21H20O6</td>
<td></td>
</tr>
<tr>
<td>Molecular weight</td>
<td>368.38 g/mol</td>
<td>150 ~ 500 g/mol</td>
</tr>
<tr>
<td>Num. heavy atoms</td>
<td>27</td>
<td></td>
</tr>
<tr>
<td>Num. arom. heavy atoms</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Fraction Csp3</td>
<td>0.14</td>
<td>0.25 ~ &lt; 1</td>
</tr>
<tr>
<td>Num. rotatable bonds</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>Num. H-bond acceptors</td>
<td>6</td>
<td>0 ~ 10</td>
</tr>
<tr>
<td>Num. H-bond donors</td>
<td>2</td>
<td>0 ~ 5</td>
</tr>
<tr>
<td>Molar Refractivity</td>
<td>102.80</td>
<td>40 ~ 130</td>
</tr>
<tr>
<td>TPSA</td>
<td>93.06 Å²</td>
<td>20 ~ 130</td>
</tr>
</tbody>
</table>

Lipophilicity

Log P<sub>o/w</sub> (XLOGP3) 3.20 - 0.7 ~ 5

Water Solubility

Log S (ESOL) -3.94 0 ~ 6

Solubility 4.22e-02 mg/ml ; 1.15e-04 mol/l

Class Soluble

Pharmacokinetics

GI absorption High

BBB permeant No

P-gp substrate No

CYP1A2 inhibitor No

CYP2C19 inhibitor No

Drug-likeness

Lipinski Yes; 0 violation

Ghose Yes

Veber Yes

Egan Yes

Muegge Yes

Bioavailability Score 0.55 >0.1

Synthetic accessibility 2.97 1 (very easy) ~ 10 (very difficult)

TPSA; topological polar surface area, GI; Gastro-Intestinal Absorption, BBB; Blood brain barrier.

3.3. Curcumin target are enriched in cancer associated signaling pathways

We conducted protein interaction networks of these identified curcumin target genes using the STRING algorithm (Figure 2A) and found that, the enriched KEGG pathways in the gene network clustering include AMPK signaling pathway, mTOR signaling pathway, PI3K-Akt signaling pathway, Insulin resistance, Endometrial cancer, Insulin signaling pathway and EGFR tyrosine kinase inhibitor resistance with strong association strength range between 1.41 to 1.99 and false discovery range of between 1.03E-10 to 8.04E-12 (Table 3). The biological process enriched in the pathway include the intracellular signal transduction, cellular response to nerve growth factor stimulus, regulation of phosphate metabolic process, apoptotic process and cellular response to nerve growth factor stimulus (Table 3). The second order clustering of the identified curcumin targets generated interactions of 34 and 182 nodes and edges.
respectively with average local clustering coefficient of 0.685 and PPI enrichment value of $1.0 \times 10^{-16}$ (Figure 2B and accompanying table).

**Table 2** Potential gene targets by curcumin

<table>
<thead>
<tr>
<th>Target</th>
<th>Gene Code</th>
<th>ChEMBL ID</th>
<th>Probability</th>
<th>Target Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microtubule-associated protein tau</td>
<td>MAPT</td>
<td>CHEMBL1293224</td>
<td>1</td>
<td>Unclassified</td>
</tr>
<tr>
<td>Toll-like receptor 9</td>
<td>TLR9</td>
<td>CHEMBL5804</td>
<td>1</td>
<td>Unclassified</td>
</tr>
<tr>
<td>Tyrosyl-DNA phosphodiesterase 1</td>
<td>TDP1</td>
<td>CHEMBL1075138</td>
<td>1</td>
<td>Enzyme</td>
</tr>
<tr>
<td>Muscle blind-like protein</td>
<td>MBNL1</td>
<td>CHEMBL1293317</td>
<td>0.77</td>
<td>Unclassified</td>
</tr>
<tr>
<td>Lactoylglutathione lyase</td>
<td>GLO1</td>
<td>CHEMBL2424</td>
<td>0.75</td>
<td>Enzyme</td>
</tr>
<tr>
<td>RAC-alpha serine/threonine-protein kinase</td>
<td>AKT1</td>
<td>CHEMBL4282</td>
<td>0.66</td>
<td>Ser_Thr Kinase</td>
</tr>
<tr>
<td>RAC-beta serine/threonine-protein kinase</td>
<td>AKT2</td>
<td>CHEMBL2431</td>
<td>0.66</td>
<td>Ser_Thr Kinase</td>
</tr>
<tr>
<td>RAC-gamma serine/threonine-protein kinase</td>
<td>AKT3</td>
<td>CHEMBL4816</td>
<td>0.66</td>
<td>Ser_Thr Kinase</td>
</tr>
<tr>
<td>Testosterone 17-beta-dehydrogenase 3</td>
<td>HSD17B3</td>
<td>CHEMBL4234</td>
<td>0.62</td>
<td>Enzyme</td>
</tr>
<tr>
<td>Estradiol 17-beta-dehydrogenase 12</td>
<td>HSD17B12</td>
<td>CHEMBL5998</td>
<td>0.62</td>
<td>Enzyme</td>
</tr>
<tr>
<td>Quinone oxidoreductase</td>
<td>CRYZ</td>
<td>CHEMBL6118</td>
<td>0.57</td>
<td>Enzyme</td>
</tr>
<tr>
<td>Gamma-secretase C-terminal fragment 59</td>
<td>APP</td>
<td>CHEMBL2487</td>
<td>0.44</td>
<td>Membrane receptor</td>
</tr>
</tbody>
</table>

**Figure 2** Curcumin targets are enriched in cancer associated signaling pathways. (A) STRING PPI direct interaction network of the identified curcumin targets. (B) STRING PPI second order clustering of the identified curcumin targets. A total of 34 and 182 nodes and edges respectively were observed with average local clustering coefficient of 0.685 and PPI enrichment value of $1.0 \times 10^{-16}$.
Table 3 Enriched KEGG and gene ontology pathways in curcumin target network

<table>
<thead>
<tr>
<th>#term ID</th>
<th>Term description</th>
<th>Observed gene count</th>
<th>Background gene count</th>
<th>Strength</th>
<th>False discovery rate</th>
<th>Matching proteins in your network (labels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>hsa04152</td>
<td>AMPK signaling pathway</td>
<td>9</td>
<td>120</td>
<td>1.79</td>
<td>2.66E-12</td>
<td>TSC2,AKT3,PIK3CA,PDPK1,MTOR,AKT2,FOXO3,PPP2CA,AKT1</td>
</tr>
<tr>
<td>hsa04150</td>
<td>mTOR signalling pathway</td>
<td>9</td>
<td>148</td>
<td>1.7</td>
<td>8.04E-12</td>
<td>TSC2,AKT3,PIK3CA,RICTOR,GSK3B,PDPK1,MTOR,AKT2,AKT1</td>
</tr>
<tr>
<td>hsa04151</td>
<td>PI3K-Akt signaling pathway</td>
<td>11</td>
<td>348</td>
<td>1.41</td>
<td>8.04E-12</td>
<td>TSC2,AKT3,PIK3CA,NOS3,GSK3B,PDPK1,MTOR,AKT2,FOXO3,PPP2CA,AKT1</td>
</tr>
<tr>
<td>hsa04931</td>
<td>Insulin resistance</td>
<td>8</td>
<td>107</td>
<td>1.78</td>
<td>2.36E-11</td>
<td>AKT3,PIK3CA,NOS3,GSK3B,PDPK1,MTOR,AKT2,AKT1</td>
</tr>
<tr>
<td>hsa05213</td>
<td>Endometrial cancer</td>
<td>7</td>
<td>58</td>
<td>1.99</td>
<td>2.65E-11</td>
<td>AKT3,PIK3CA,GSK3B,PDPK1,MTOR,AKT2,AKT1</td>
</tr>
<tr>
<td>hsa04140</td>
<td>Autophagy - animal</td>
<td>8</td>
<td>125</td>
<td>1.72</td>
<td>4.41E-11</td>
<td>TSC2,AKT3,PIK3CA,PDPK1,MTOR,AKT2,PPP2CA,AKT1</td>
</tr>
<tr>
<td>hsa04910</td>
<td>Insulin signaling pathway</td>
<td>8</td>
<td>134</td>
<td>1.69</td>
<td>6.56E-11</td>
<td>TSC2,AKT3,PIK3CA,GSK3B,PDPK1,MTOR,AKT2,AKT1</td>
</tr>
<tr>
<td>hsa01521</td>
<td>EGFR tyrosine kinase inhibitor resistance</td>
<td>7</td>
<td>78</td>
<td>1.86</td>
<td>1.03E-10</td>
<td>AKT3,PIK3CA,GSK3B,MTOR,AKT2,FOXO3,AKT1</td>
</tr>
<tr>
<td>GO:0035556</td>
<td>intracellular signal transduction</td>
<td>14</td>
<td>1528</td>
<td>0.87</td>
<td>6.65E-07</td>
<td>TSC2,AKT3,PIK3CA,APP,RICTOR,NOS3,GSK3B,PDPK1,TLR9,MTOR,AKT2,FOXO3,PPP2CA,AKT1</td>
</tr>
<tr>
<td>GO:0006915</td>
<td>apoptotic process</td>
<td>11</td>
<td>915</td>
<td>0.99</td>
<td>1.47E-06</td>
<td>TSC2,PIK3CA,APP,GSK3B,PDPK1,MTOR,AKT2,FOXO3,PPP2CA,AKT1,APBB1</td>
</tr>
<tr>
<td>GO:1999090</td>
<td>cellular response to nerve growth factor stimulus</td>
<td>5</td>
<td>44</td>
<td>1.97</td>
<td>1.47E-06</td>
<td>APP,MAPT,PDPK1,FOXO3,AKT1</td>
</tr>
<tr>
<td>GO:0080134</td>
<td>regulation of response to stress</td>
<td>12</td>
<td>1299</td>
<td>0.88</td>
<td>1.77E-06</td>
<td>AKT3,APP,RICTOR,NOS3,GSK3B,MAPT,PDPK1,TLR9,MTOR,FOXO3,AKT1,APBB1</td>
</tr>
<tr>
<td>GO:0033674</td>
<td>positive regulation of kinase activity</td>
<td>9</td>
<td>553</td>
<td>1.12</td>
<td>1.84E-06</td>
<td>PIK3CA,APP,RICTOR,MAPT,PDPK1,TLR9,MTOR,PPP2CA,AKT1</td>
</tr>
</tbody>
</table>
3.4. Curcumin targets are over-expressed in multiple cancers and are predictors of poor prognosis

We conducted a gene differential expression analysis of the three top ranked curcumin target genes (Microtubule-associated protein tau (MAPT), Toll-like receptor 9 (TLR9) and Tyrosyl-DNA phosphodiesterase 1 (TDP1)), between cancer patient and normal patient and found that MAPT, TLR and TDP1 are overexpressed (p < 0.05) in multiple panels of cancer including; breast cancer, colon cancer, head and neck cancer, Kidney cancer, liver cancer, PAAD, melanoma, STAD, THCA and UECE (figure 3). Furthermore, we also found that higher RNA expression profile of MAPT, TLR9 and TDP1 in all TCGA and the GTEx datasets correlate with low overall survival and disease free survival of cancer patients. (Figure 4)

**Figure 4** Box plots showing the gene expression levels of Microtubule-associated protein tau (MAPT), Toll-like receptor 9 (TLR9) and Tyrosyl-DNA phosphodiesterase 1 (TDP1) in different human between panels of human cancer patient and health patient according to TCGA database. Blue label (normal tissue) and red label represent tumor sample. The statistical significance of the differential expressed gene was evaluated using the Wilcoxon test.
4. Conclusion

In conclusion, this study identified 12 genes as a druggable candidate for curcumin. The genes identified were enriched in cancer and drug resistance associated signaling pathway. Interestingly, the top 3 identified genes; Microtubule-associated protein tau (MAPT), Toll-like receptor 9 (TLR9) and Tyrosyl-DNA phosphodiesterase 1 (TDP1) were observed to be over expressed in multiple cancer cohorts and were associated with poor prognoses of clinical cancer cohorts. Curcumin has good physicochemical, bioavailability and ADMET properties. Importantly, it met the Lipinski’s Rule of 5 for drug likeness and thus worthy of further in vitro and in vivo confirmation studies on the targets identified in this study.

Compliance with ethical standards

Acknowledgments

We acknowledged the Swiss Institute of Bioinformatics for designing most of the algorithm used in this study.

Disclosure of conflict of interest

The authors declare no conflicts of interest.

Author contributions

All authors participated in data collection, visualization, analysis and writing the manuscript. All authors read and approved the final manuscript.
References


