Bioorganic & Medicinal Chemistry 20 (2012) 1251–1258

Contents lists available at SciVerse ScienceDirect

Bioorganic & Medicinal Chemistry

journal homepage: www.elsevier.com/locate/bmc

Construction of an Indonesian herbal constituents database and its use in Random Forest modelling in a search for inhibitors of aldose reductase

Sadaf Naeem, Peter Hylands, David Barlow *

Institute of Pharmaceutical Science, Franklin-Wilkins Building, King's College London, 150 Stamford Street, London SE1 9NH, United Kingdom

ARTICLE INFO

Article history:
Received 30 August 2011
Revised 12 December 2011
Accepted 17 December 2011
Available online 30 December 2011

Keywords:
Aldose reductase
Random Forest modelling
Indonesian herbal medicine

ABSTRACT

Data on phytochemical constituents of plants commonly used in traditional Indonesian medicine have been compiled as a computer database. This database (the Indonesian Herbal constituents database, IHD) currently contains details on ~1,000 compounds found in 33 different plants. For each entry, the IHD gives details of chemical structure, trivial and systematic name, CAS registry number, pharmacology (where known), toxicology (LD_{50}), botanical species, the part(s) of the plant(s) where the compounds are found, typical dosage(s) and reference(s). A second database has been also been compiled for plant-derived compounds with known activity against the enzyme, aldose reductase (AR). This database (the aldose reductase inhibitors database, ARID) contains the same details as the IHD, and currently comprises information on 120 different AR inhibitors. Virtual screening of all compounds in the IHD has been performed using Random Forest (RF) modelling, in a search for novel leads active against AR—to provide for new forms of symptomatic relief in diabetic patients. For the RF modelling, a set of simple 2D chemical descriptors were employed to classify all compounds in the combined ARID and IHD databases as either active or inactive as AR inhibitors. The resulting RF models (which gave misclassification rates of 21%) were used to identify putative new AR inhibitors in the IHD, with such compounds being identified as those giving RF scores >0.5 (in each of the three different RF models developed). In vitro assays were subsequently performed for four of the compounds obtained as hits in this in silico screening, to determine their inhibitory activity against human recombinant AR. The two compounds having the highest RF scores (prunetin and ononin) were shown to have the highest activities experimentally (giving ~58% and ~52% inhibition at a concentration of 15 μM, respectively), while the compounds with lowest RF scores (vanillic acid and cinnamic acid) showed the lowest activities experimentally (giving ~29% and ~44% inhibition at a concentration of 15 μM, respectively). These simple virtual screening studies were thus helpful in identifying novel inhibitors of AR, but yielded compounds with only very modest (micromolar) potency.

1. Introduction

Plants have been utilized as medicines for thousand of years[1] and the great civilisations of the ancient Chinese, Indians and North Africans provide written evidence of the utilization of plants for the treatment of a wide variety of diseases.[2] Since the sequencing of the human genome, thousands of new molecular targets have been identified as being involved in various diseases[3] and known compounds from medicinal plants show promising and possibly selective activity towards these targets.[4] Indeed, there are numerous compounds quite recently isolated from traditionally-used medicinal plants that have been developed as modern pharmaceuticals (e.g., huperzine[5,6] and paclitaxel/Taxol®.[7–11]

A recent analysis by Fabricant and Farnsworth[12] showed that the uses of 80% of 122 plant derived drugs were related to their original ethno-pharmacological purposes. It was also discovered that these compounds were derived from just 94 different species of plants.[13]

Because these compounds are derived from only 94 species of plants, and a conservative estimate of the number of higher plants occurring on the planet is 250,000[12] it is clear that a large part of nature's chemical diversity remains to be exploited, with innumerable drugs remaining to be discovered from these plants.

In Indonesia where, even today, herbal medicine is popular, more than 1,300 species are known as medicinal plants called '*jamu*'.[14] Nowadays, *jamu* is being developed from traditional handling to industrial (large scale) production. World wide, however, *jamu* is less well known than Traditional Chinese Medicine (TCM), Japanese Kampo or Indian Ayurveda. Moreover, since the various biological activities ascribed to *jamu* are only rather poorly founded on empirical data, there is a clear need for more scientific research to establish efficacy and ensure safety. A scientific approach is essential to further develop the rational use of *jamu*.

* Corresponding author. Tel.: +44 (0) 207 848 4827; fax: +44 (0) 207 848 4800.
E-mail address: dave.Barlow@kcl.ac.uk (D. Barlow).

0968-0896/$ - see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.bmc.2011.12.033

The drug discovery process is such a money and time consuming process that it becomes more and more difficult for the industry to develop new drugs, and there is a growing emphasis on the development and use of structural (computer) database searches. Among the various databases available, those holding information on compounds derived from natural products attract particular interest, since natural products are widely seen as an abundant source of candidate drugs.[15–18]

In the work reported here, a database of the chemical constituents from Traditional Indonesian medicinal plants has been assembled, and ligand-based virtual screening of these compounds by Random Forest modelling has been performed in a search for potential inhibitors of aldose reductase-the first and rate-limiting enzyme of the polyol pathway that catalyses the conversion of glucose to sorbitol. It is the activation of this enzyme in diabetic patients that leads to complications including retinopathy, neuropathy and cataract. Inhibition of the enzyme may therefore prevent or delay the onset of these complications.

As a means to assess the reliability of the RF screening, in vitro assays of inhibitory activity against human recombinant aldose reductase (h-rAR) were performed for selected hits.

2. Materials and methods

2.1. Data sets

Two chemical databases were constructed using ChemDBSoft 3.1 (TimTec Inc, USA): the first contains information on phytochemical compounds with known activity against the aldose reductase (AR) enzyme (the aldose reductase inhibitors database, ARID), compiled from Duke's Phytochemical and Ethnobotanical Database[19]; the second, provides details of the known chemical constituents of some of the plants important in Indonesian traditional medicine, '*jamu*' (the Indonesian Herbal constituents database, IHD). The plants used in traditional Indonesian medicine were taken from Roosita,[14] Usia,[20] Grosvenor[21] and Batubara.[22] The details of all IHD and ARID chemicals were compiled from Duke's Phytochemical and Ethnobotanical Database,[19] the NCBI PubChem database[23] and the Dictionary of Natural Products.[24]

In both the ARID and IHD databases, the details of the chemical constituents included structure, systematic name, CAS registry number, pharmacological and toxicological information, IC_{50} (where known), plant part in which the compound is found, and chemical references.

2.2. Random Forest modelling

Random Forest (RF) methodology (as implemented within Statistica version 9.1, Statsoft Ltd, Bedford, UK) was used to develop a classification model for predicting AR inhibitors. RF, one of the best known forms of multiple decision tress, uses a small set of active compounds to search a much larger data set for new candidates. In the studies reported here, RF was used to search for IHD compounds that might be active against AR. Compounds from the ARID were added to the target IHD, the former constituents providing the minority class, the latter, the majority class.

Table 1
Known aldose reductase inhibitors included as entries in ARID

S. no	Compound	S. no	Compound	S. no	Compound
1	Glaucine	39	Eupatilin	76	*p*-Coumaric acid
2	3,3',4-Tri-*o*-methyl ellagic acid	40	Forsythoside B	77	Pectolinarigenin
3	3,4-Dimethoxy cinnamic acid	41	Fraxin	78	Pectolinarin
4	7-Methylcapillarisin	42	Fraxinol	79	Protopine
5	Acacetin	43	Gamma terpinen	80	Prunasin
6	Acteoside	44	Genistein	81	Quercetagetin
7	Allo-cymene	45	Gossypol	82	Quercetin
8	Allocryptopine	46	Haematoxylin	83	Quercimeritrin
9	Alpha terpinene	47	Herniarin	84	Quercitrin
10	Alpha terpineol	48	Hymenoxin	85	Resorcinol
11	Amentoflavone	49	Hyperin	86	Reynoutrin
12	Apiin	50	Indole	87	Rhamnetin
13	Ascorbic acid	51	Isofraxidin	88	Rhamnocitrin
14	Astragalin	52	Isoliquiritigenin	89	Rhoifolin
15	Avicularin	53	Isoliquiritin	90	Rosmarinic acid
16	Axillarin	54	Isoquercitrin	91	Rutin
17	Baicalein	55	Isoscopoletin	92	Salicylic acid
18	Baicalin	56	Kaempferol	93	Salvianolic acid A
19	Caffeic acid	57	Kaempferol-3-rhamnoside	94	Scoparone
20	Capillarisin	58	Kuwanon C	95	Scopoletin
21	Caryophyllene	59	Licorisidin	96	Scutellarein
22	Chlorogenic acid	60	Licuraside	97	Sideritoflavone
23	Chrysoeriol	61	Linarin	98	Skullcapflavone II
24	Chrysosplenol	62	Lonicerin	99	Sophoricoside
25	Cinnamic acid	63	Luteolin	100	Spiraeoside
26	Cirsilineol	64	Luteolon-7-glucoside	101	Sugiol
27	Cirsiliol	65	Luteolin-7-glucuronide	102	Tanshinone I
28	Citrinin	66	Morin	103	Tanshinone II
29	Cosmosiin	67	Myrcene	104	Taxifolin
30	Coumarin	68	Myricitrin	105	Thymonin
31	Cryptotanshinone	69	Naringenin	106	Trifolin
32	Cynaroside	70	Naringin	107	Umbelliferone
33	Daucosterol	71	Nelumbdside	108	Vanillic acid
34	Dehydrocorydaline	72	Nepetin	109	Verbascoside
35	Delta cadinene	73	Nerolidol	110	Vitexin
36	Diosmetin	74	Nevadensin	111	Wogonin
37	Echinatin	75	Oroxylin A	112	Xanthomicrol
38	Ellagic acid		—		—

The RF methodology adopted for construction of the classification model was as described by Ehrman et al.[25]:

(a) For each tree in the ensemble, a bootstrap sample was taken from the minority class, with the same number of cases then drawn randomly from the majority class.
(b) A classification tree was built to maximum size without pruning. Classification was made according to a simple binary scheme—so that the various phytochemicals were predicted to be AR inhibitors or IHD constituents—on the basis of their computed molecular descriptors (see below). The tree building employed the CART algorithm using the Gini splitting criterion[26] with the modification that a random subset of molecular descriptors (D_{rand}) was searched at each node in the tree as opposed to the full set.
(c) Steps (a) and (b) were repeated for every tree built. The final predictions of class membership were made on the basis of the consensus of all predictions from the ensemble of trees.

'Out of Bag' (OOB) cross validation was employed so that for each tree generated approximately 33% of compounds were randomly excluded and treated as an independent test sample. To measure the generalization error for each tree, the OOB sample was run through that tree and the error rate of prediction

Table 2
List of Indonesian herbs present in IHD, their local name, part of plant used, therapeutic application(s) and the number of identified chemical constituents

Indonesian herb	Local name	Part used	Therapeutic application	No. of compounds	Reference
Abelmoschus moschatus	Kapas hantu	Leaf	Nausea	16	21
Abrus precatorius	Saga	Leaf	Postpartum remedy, cough, urinary disease	14	14
Ageratum conyzoides	Babadotan	Leaf	Postpartum remedy, lung disease, lack of appetite	50	14
Aleurites moluccana	Kemiri	Fruits, Leaves	Headache, fever, inflammation, gonorrhoea, the lowering of cholesterol	7	—
Allium cepa	Bawang merah	Tuber	Fever, postpartum remedy, stomachache, measles	99	14
Allium sativum	Bawang puith	Tuber	Fever, measles	82	14
Alpinia galanga	Laos/Lengkuas	Rhizome/ Tuber	Stomachic, anorexia, muscle pain	15	14,20
Alstonia scholaris	Pulai	Leaf, Bark	Beriberi, syphilis, diabetes, fever anorexia, nephritis, diabetes, malaria, hypertension	3	20
Amomum compactum	Kapulaga	RhizomeFruit	Cough, tonsillitis, menstrual disorder, colic, influenza, gastritis	18	20
Andrographis paniculata	Sambiloto	LeafAerial part	Fever, diabetes, gonorrhea, syphilis, tonsillitis	4	14,20
Annona muricata	Nangka walandaSirsak	Leaf	Dermatitis	28	14
Artemisia vulgaris	Siang	Leaf	Postpartum remedy	24	14
Canna indica	Ganyong	Leaf	Postpartum remedy	2	14
Carica papaya	Gedang gandul	Leaf	Fever, postpartum remedy, muscle pain	74	14
Catharanthus roseus	Tapak dara	Aerial part	Diabetes, cancer, malaria, hypertension	55	20
Ceiba pentandra	Kapas Randu	Leaf	Dermatitis	17	14
Cocos nucifera	Kelapa	Stem, Fruits	Fever, hepatitis	37	14
Curcuma xanthorrhiza	Temu lawak	TuberRhizome	Postpartum remedy, muscle pain, hepatitis, anticonvulsant, hemorrhoids, malaria, diarrhoea, anorexia, gastritis, anemia	14	14,20
Cymbopogon nardus	Sere	Aerial part	Diaphoretic, body warming	41	20
Durio zibethinus Murray	Duren	Bark	Dysentery	0	14
Equisetum debile	—	—	—	6	—
Eryngium foetidum	Walang	Leaf	Fever	9	14
Foeniculum vulgare	Adhas	Seed	Albuminurea, insomnia, menstrual disorder	95	20
Glycyrrhiza glabra	Kayu legi	Stem	Hepatitis, tonic	194	20
Melalecla leucadendron	Kayu putih	Leaf	Vertigo, anticonvulsant, toothache	3	20
Piper nigrum	Lada, Merica putih	Fruit, Leaf	Carminative, hypertension, dyspnea, toothache, muscle pain	109	14,20
Piper cubeba	Kemukus	Fruit	Dysentery, gonorrhea	41	20
Punica granatum	Delima puith	Fruit	Leucorrhea, constipation, diarrhea, anorexia, gastritis, anemia, jaundice, dysentery	46	20
Rheum palmatum	Klembak	Root	Tonic, Stomach-ache	16	20
Santalum album	Kayu cendana	Wood	Dysentery, asthma, fever, gonorrhoea	35	20
Syzygium aromaticum	Cengkeh	Flower	Cold, cough	63	20
Tamarindus indica	Asam jawa	Fruit, Leaf	Fever, laxative Postpartum-medication, measles	40	14,20
Terminalia catappa	Ketapang	—	Dysentry, small pox	16	22

measured. The error rates for all trees were then averaged to give the overall generalization error for the entire forest.

To identify the potential inhibitors of AR within the IHD, five different RF models were generated, with each one using a different set of molecular descriptors. There were 112 (AR) active compounds used in the training set together with 1,242 IHD compounds.

2.3. Molecular descriptors used in RF modelling

Five sets of molecular descriptors were calculated for each entry in both data sets using DRAGON Professional version 5.5 (2007), Talete SRL, Italy, that are: Constitutional descriptors; Two dimensional (2D) autocorrelation descriptors[27–29]; Three dimensional (3D)-MoRSE descriptors[30]; Radial Distribution Function (RDF) descriptors[31] and Randic molecular profiles.[32,33]

3. In vitro assay

3.1. Materials

Human recombinant aldose reductase (h-rAR) enzyme was purchased from Source BioScience Lifesciences PLC (Nottingham, UK; Cat No. ABC2942), and β-NADPH coenzyme from Calbiochem-Novabiochem Ltd (Nottingham, UK; Cat.No. 481973). Dipotassium hydrogen phosphate (K_2HPO_4) and mono-potassium dihydrogen phosphate (KH_2PO_4) were purchased from BDH (Chester, UK) and dimethyl sulfoxide (DMSO) from Fisher Scientific (Loughborough,

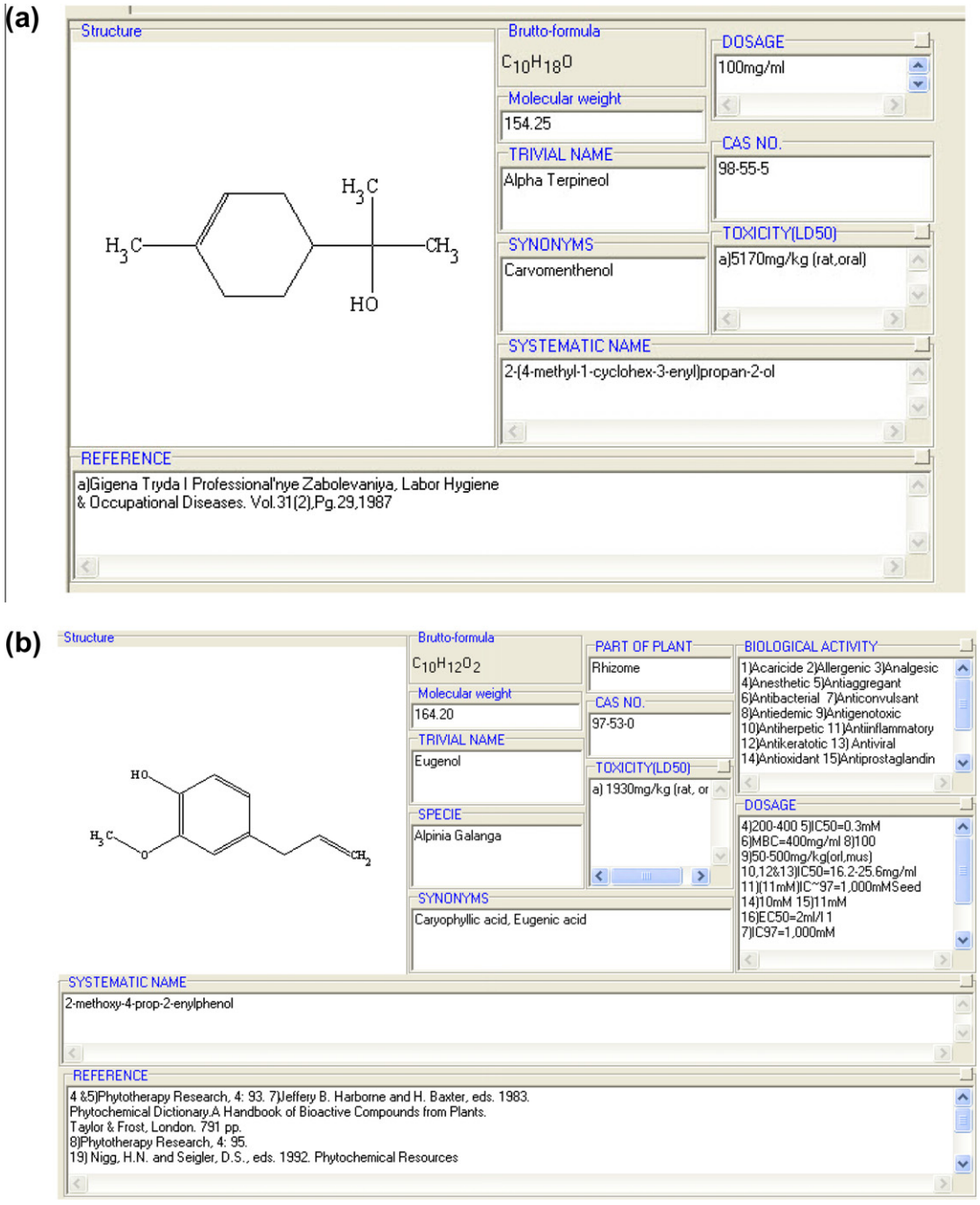

Figure 1. Screen shot of the ARID entry for α terpineol (a) and the IHD entry for eugenol (b).

Table 3
Percent accuracy of RF predictions for compounds (38) common in IHD and ARID

Descriptor set	Predicted active	Predicted inactive	Percent (%) accuracy
Constitutional	31	7	81.6
RDF	31	7	81.6
3D-MoRSE	31	7	81.6
Randic	31	7	81.6
2D autocorrelation	27	11	71.1

Table 4
Top 20 hits from RF modelling and the herbs from IHD in which these compounds are found

Top 20 RF hits	Herbs from IHD in which these compounds are found
Kumatakenin	*Glycyrrhiza glabra*
Rhamnetin	*Syzygium aromaticum*
Kaempferide	*Alpinia galanga*
Isoquercitrin	*Allium cepaAllium sativumFoeniculum vulgareArtemisia vulgarisAbelmoschus moschatus*
Hyperin	*Rheum palmatum*
Quercitrin	*Rheum palmatum*
Kampferol-3-o-glucoside	*Abelmoschus moschatus*
Astragalin	*Glycyrrhiza glabra*
Spiraeoside	*Allium cepa*
Prunetin	*Glycyrrhiza glabra*
Quercetin-3-glucoside	*Foeniculum vulgareArtemisia vulgaris*
Formononetin	*Glycyrrhiza glabra*
Ononin	*Glycyrrhiza glabra*
Kaempferol	*Allium cepaAllium sativumFoeniculum vulgareAbelmoschus moschatusAgeratum conyzoidesCeiba pentandraGlycyrrhiza glabraSyzygium aromaticumTerminalia catappa*
Isolicoflavonol	*Glycyrrhiza glabra*
Licoflavonol	*Glycyrrhiza glabra*
Licoisoflavone A	*Glycyrrhiza glabra*
Cyanidin-3-glucoside	*Terminalia catappaPunica granatum*
Pelargonidin-3-glucoside	*Punica granatum*
Isoliquiritin	*Glycyrrhiza glabra*

UK). D-glyceraldehyde, tanshinone IIA (98% purity), prunetin (>98% purity), ononin (>99% purity), cinnamic acid (>97% purity) and vanillic acid (>97% purity) were obtained from Sigma–Aldrich Company Ltd (Gillingham, UK). All chemicals were used as received.

Aldose reductase enzyme solution (received as 50 μg/50 μL) was diluted with Tris HCl buffer (20 mM; pH 8) to give a stock solution of 500 nM. D-glyceraldehyde (the substrate for aldose reductase) was prepared as a 20 mM stock in potassium phosphate buffer (100 mM; pH 7). β-NADPH (the co-enzyme for aldose reductase) was prepared as a 1 mM stock in potassium phosphate buffer (100 mM; pH 7). (Solutions were prepared fresh for each experiment, and were stored on ice until used). Tanshinone IIA was prepared as a 200 μM stock in DMSO. From this stock solution a range of different concentrations from 10 to 150 μM were prepared by diluting with potassium phosphate buffer (100 mM; pH 7) to obtain the required concentrations. All test compounds (prunetin, ononin, cinnamic acid and vanillic acid) were each prepared as 300–500 μM stock solution by dissolving in 5–10 mL DMSO and these then diluted with potassium phosphate buffer (100 mM; pH 7) to obtain 150 μM solutions for use.

3.2. Aldose reductase assay

The assay for AR activity was performed in a 96-well plate, essentially as described by Mylari et al.[34] Initiation of the reaction (with substrate) was preceded by a 10 min pre-incubation at 37 °C of potassium phosphate buffer (155 μL; 100 mM; pH 7), containing β-NADPH (25 μL; 125 μM) and human recombinant aldose reductase (20 μL; 50 nM) with 25 μL standard compound/test compound solution. The reaction was initiated by addition of 25 μL D-glyceraldehyde (20 mM). The rate of decrease in OD_{340} was monitored by means of a Spectra Max 190 Plate Reader, at 10 s intervals for a total of 50 min. The in-well concentration of glyceraldehyde was 2 mM. All the test compounds were assayed at in-well concentrations of 15 μM, and tanshinone IIA with concentrations ranging from 1 to 20 μM. In all cases the in-well DMSO levels were 1%.

The activity of the test compounds was computed as:

$$\text{Inhibitory activity}(\%) = (\Delta OD_{control} - \Delta OD_{test})/(\Delta OD_{control}) \times 100$$

in which ΔOD_{test} and $\Delta OD_{control}$ are the rates of change of the absorbances measured at 340 nm for the inhibited and uninhibited enzyme samples, respectively. All test compounds were assayed in triplicate and the uninhibited enzyme assayed with $n = 15$.

4. Results and discussion

Two ChemDBSoft databases have been constructed: an Aldose Reductase Inhibitors Database (ARID) and an Indonesian Herbal constituents Database (IHD). ARID contains information on 112 compounds from different herbs that have been shown to inhibit the enzyme aldose reductase (Table 1), and IHD contains details of 1,242 compounds from 33 different plants used in Indonesian traditional medicine. The numbers of compounds from each of the IHD herbs together with the source references used are presented in Table 2. Sample screen shots of the ChemDBSoft Databases for one ARID entry (alpha terpineol) and one IHD entry (eugenol) are shown in Figure 1.

Ehrman et al.[25] have previously shown that simple 1D and/or 2D molecular descriptors are very often perfectly adequate for predicting the biological properties of molecules. In the studies conducted here, therefore, RF models for prediction of AR inhibitors were explored using only a selection of five simple molecular descriptors. Despite the highly unbalanced ratio between active compounds (112) and the target dataset (1,242) the RF modelling results for all the five descriptor sets, proved satisfactory. The misclassification rates for the various RF models were 85.7% for the models constructed using the Randic and 2D-autocorrelation function descriptors, 81.3% for the models constructed using the

Table 5
Predicted RF score, IC_{50} (μM) and pIC_{50} for known aldose reductase inhibitors included in ARID and IHD

Compounds	Predicted RF score	IC_{50} (μM)
Rhamnetin	0.97	2.7
Isoquercitrin	0.96	4.5
Hyperin	0.96	3
Quercitrin	0.95	2
Astragalin	0.95	60.5
Spiraeoside	0.95	95.31
Kaempferol	0.91	10
Isoliquiritin	0.91	0.72
Vitexin	0.90	566.66
Genistein	0.88	20
Quercetin	0.86	2.2
Naringenin	0.86	28.87
Avicularin	0.85	19.05
Isoliquiritigenin	0.82	0.32
Licuraside	0.80	0.56
Chlorogenic acid	0.70	1.8
Ellagic acid	0.69	0.2
Herniarin	0.66	131

RDF and 3D-Morse descriptors, and 90.1% for the model constructed using the simple constitutional descriptors.

For the 38 compounds in the IHD which are also present in ARID and already known to be active against AR, the accuracy of the RF predictions made using the five different RF models is $\geqslant$70%. Given that the percent accuracy of the constitutional, RDF and 3D-MoRSE models is equally good for the known AR active compounds (81.6%, Table 3), the final ARI predictions were made using these three models.

The compounds from the IHD predicted to be active against AR by three of the RF models; Constitutional, RDF and 3D-MoRSE and the average of the three RF scores are presented in Table S1 in Supplementary data. Compounds were classified as active if they achieved a score of 0.50 or above in all three RF models. They were thus classified as active by $\geqslant$50% votes from the full ensemble of trees in each RF model.

The activities predicted in the RF screening of the IHD compounds show that approximately 55% of the hits obtained are flavonoids or flavonoid derivatives (see Table S1 in Supplementary data) and this is to be expected because of the high number of flavonoids and flavonoid-related compounds in the ARID training set (Table 1). These findings accord with those reported for TCM compounds by Ehrman et al.[25], where most of the compounds predicted as ARIs were likewise found to be flavonoids (and for the same reason). The number of compounds classified as active by RF modelling for each IHD herbs are presented in Table S2 in Supplementary data.

Among the 33 herbs in the IHD, *Glycyrrhiza glabra*, has the highest number of hits predicted active against AR by the RF modelling, followed by *Rheum palmatum* and *Foeniculum vulgare* (Table S2 in Supplementary data). It is encouraging to note, therefore, that there are already a number of reports of compounds derived from

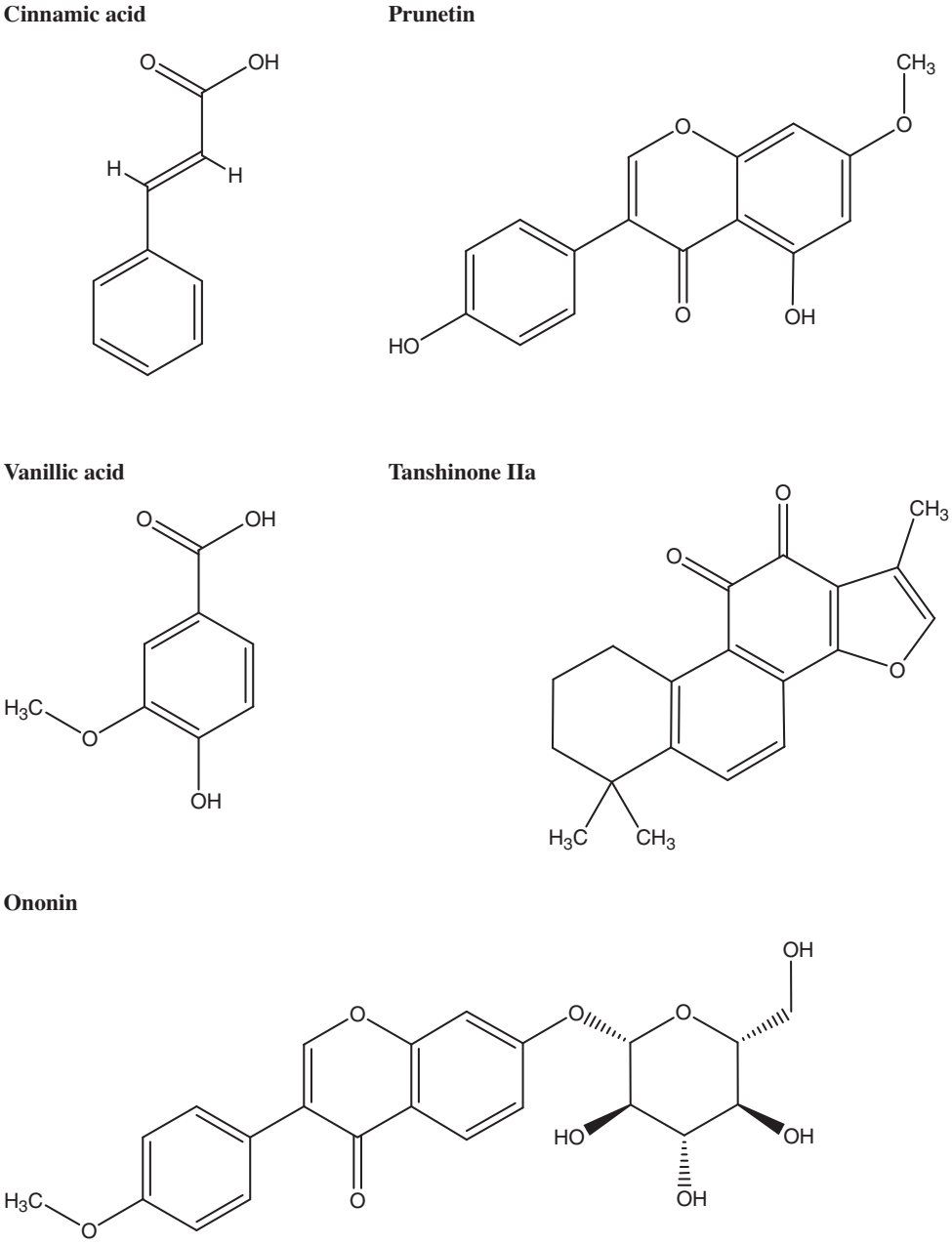

Figure 2. Chemical structures of the known aldose reductase inhibitor, tanshinone IIa, and the four test compounds identified as putative aldose reductase inhibitors by RF modelling: prunetin, vanillic acid, cinnamic acid and ononin.

Licorice species (*Glycyrrhiza* sp.) that have been identified as potentially useful for treating diabetes and diabetes-related complications, and constituents of these plants that show inhibitory activity against AR.[35,36]

Rheum palmatum has also been shown to have anti-diabetic activity, albeit through inhibition of the enzyme α-glucosidase.[37]This observation taken together with the predictions here of its constituents that may be active against AR, would indicate that this herb might be quite efficient not only in the treatment of diabetes but may also be of use in preventing the secondary complications in diabetic patients.

Although only two of the constituents from *Alpinia galanga* are predicted here by RF modelling to be active against AR, it may be noted that this herb too has been investigated in the context of diabetes treatment, and has been shown to exhibit hypoglycaemic activity in rabbits.[38]

It is observed that the herbs from IHD that are used traditionally in Indonesia for the treatment of diabetes, are *Alstonia scholaris*, *Andrographis paniculata* and *Catharanthus roseus* (Table 2). For these three herbs there were no ARI hits found in the RF screening. On the basis of these studies, however, it is impossible to judge whether this absence of hits is significant or not: it may be because of limitations in the screening processes or because these herbs have constituents helpful in controlling blood glucose levels but none which are active/useful in preventing diabetic complications. The only RF hits active against AR found in *Catharanthus roseus* are quercetin and malvidin, and these are common flavonoids which frequently feature as hits against many different targets (*cf*., Ehrman et al.[25]

The IHD herbs in which the top 20 RF ARI hits are found show similar results (Table 4), with most of the top hits obtained from RF being constituents of *Glycyrrhiza glabra* and *Rheum palmatum*. It is also observed here, that isoquercetrin and kaempferol, among the top 20 RF hits, are common flavonoids and found in a variety of plants.

Comparison of the RF scores and IC_{50} values for the known ARIs included in ARID (Table 5) shows that these compounds have mean RF scores >0.5, and generally have inhibitory activities in the range 120–0.2 μM. It should be noted, however, that the RF scores are not correlated with IC_{50} because the RF modelling allows only that compounds are classified as 'active' or 'inactive' and provides no information as to their relative activities.

Four compounds from the list (Table 2)—prunetin, ononin, cinnamic acid and vanillic acid (Fig. 2) were selected for in vitro assay of AR inhibitory activity. These compounds have RF scores ranging from 0.94 (prunetin) down to 0.63 (vanillic acid). On the basis of the findings for known ARIs (Table 5), therefore, it might be expected that these compounds would have inhibitory activities in the micromolar range.

As a positive control for these in vitro tests, tanshinone IIA (Fig. 2) was also assayed for ARI activity. Tanshinone IIA has a reported IC_{50} of 1.14 μM against rat lens aldose reductase (rAR).[39]The percent inhibition results obtained for this compound against human recombinant aldose reductase (h-rAR) are presented in Table 6. Solubility constraints (imposed by the need to maintain a low concentration of DMSO in the test solution) precluded assays involving >20 μM concentrations of tanshinone IIA, and so it was not possible to obtain a reliable estimate of IC_{50} from the incomplete dose–response curve. It is nevertheless clear from the data obtained (Table 6) that the IC_{50} for this compound must lie in the region of 7.5 μM to 10 μM—these concentrations giving ~45% and ~54% inhibition, respectively. (The curve fit obtained for these data using Origin 7.5 (OriginLab, MA, USA) is presented in Figure S1 in Supplementary data, and yields an IC_{50} of 8.69 μM.) It thus seems that tanshinone IIA is a rather poorer inhibitor of h-rAR than of rAR (IC_{50} = 1.14 μM;[40]). This is in keeping, however, with the findings reported by Lee et al.[36] which show that compounds generally exhibit 3–10 times stronger inhibitory activity against rAR compared with h-rAR.

Table 6
Percent inhibition of h-rAR due to Tanshinone IIA at concentrations ranging from 20 μM to 1.0 μM

Tanshinone IIA (μM)	Percent (%) inhibition			Average percent (%) inhibition
20	76.7	74.4	74.4	75.2 ± 1.3
15	57.7	57.7	—	57.7 ± 0.0
10	53.8	55.7	51.7	53.7 ± 2.0
7.5	47.4	42.1	—	44.8 ± 3.7
5.0	30.8	31.5	35.6	32.6 ± 2.6
2.5	30.8	23.1	—	27.0 ± 5.4
1.0	15.4	20	—	17.7 ± 3.3

Table 7
Percent inhibition of h-rAR by selected hits from RF modelling

Test compounds	Percent (%) inhibition (15 μM)			Average percent (%) inhibition
Prunetin	57.1	57.5	60.6	58.4 ± 1.92
Ononin	52.3	52.3	52.3	52.3
Cinnamic acid	39.4	42.4	49.1	43.6 ± 4.97
Vanillic acid	30	30	28.3	29.4 ± 0.98

The experimentally measured percent inhibition for each of the four selected RF hits are presented in Table 7. These compounds thus have activities in the micromolar range and this is entirely consistent with the expectations based on inspection of the RF scores and IC_{50}s for the known ARIs (Table 5). While the hits identified through the RF modelling cannot, therefore, be considered as providing leads worthy of further development as anti-diabetic drugs, it is encouraging to note that the modelling was nevertheless successful in identifying these compounds as ARIs.

Acknowledgments

This study was funded with support from the University of Karachi, the Higher Education Commission (HEC) Pakistan, the Charles Wallace Trust and the Dr. Wali Mohammad Trust (High Commission of Pakistan).

Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.bmc.2011.12.033. These data include MOL files and InChiKeys of the most important compounds described in this article.

References and notes

1. Samuelsson, G. *A textbook of pharmacognosy*, fifth Swedish; Pharmaceutical Press: Stockholm, 2004.
2. Phillipson, J. D. *Phytochemistry* **2001**, *56*, 237.
3. Kramer, R.; Cohen, D. *Nat. Rev. Drug Discovery* **2004**, *3*, 965.
4. Balunas, M. J.; Kinghorn, A. D. *Life Sci.* **2005**, *78*, 431.
5. Desilets, A. R.; Gickas, J. J.; Dunican, K. C. *Ann. Pharmacother.* **2009**, *43*, 514.
6. Little, J. T.; Walsh, S.; Aisen, P. S. *Expert Opin. Invest. Drugs* **2008**, *17*, 209.
7. Horwitz, S. B. *J. Nat. Prod.* **2004**, *67*, 136.
8. Cragg, G. M. *Med. Res. Rev.* **1998**, *18*, 315.
9. Wall, M. E. *Med. Res. Rev.* **1998**, *18*, 299.
10. Cragg, G. M.; Newman, D. J. *J. Nat. Prod.* **2004**, *67*, 232.
11. Oberlies, N. H.; Kroll, D. J. *J. Nat. Prod.* **2004**, *67*, 129.
12. Fabricant, D. S.; Farnsworth, N. R. *Environ. Health Perspect.* **2001**, *109*, 69.
13. Farnsworth, N. R.; Akerele, O.; Bingel, A. S.; Soejarto, D. D.; Guo, Z. *Bull WHO* **1985**, *63*, 965.
14. Roosita, K.; Kusharto, C. M.; Sekiyama, M.; Fachrurozi, Y.; Ohtsuka, R. *J. Ethnopharmacol.* **2008**, *115*, 72.

15. Wang, S. M.; Zaharevitz, D. W.; Sharma, R.; Marquez, V. E.; Lewin, N. E.; Du, L.; Blumberg, P. M.; Milne, G. W. A. *J. Med. Chem.* **1994**, *37*, 4479.
16. Wang, S. M.; Milne, G. W. A.; Yan, X. J.; Posey, I. J.; Nicklaus, M. C.; Graham, L.; Rice, W. G. *J. Med. Chem.* **1996**, *39*, 2047.
17. Neamati, N.; Hong, H. X.; Mazumder, A.; Wang, S. M.; Sunder, S.; Nicklaus, M. C.; Milne, G. W. A.; Proksa, B.; Pommier, Y. *J. Med. Chem.* **1997**, *40*, 942.
18. Kiyama, R.; Honma, T.; Hayashi, K.; Ogawa, M.; Hara, M.; Fujimoto, M.; Fujishita, T. *J. Med. Chem.* **1995**, *38*, 2728.
19. Duke's Phytochemical and Ethnobotanical Database: accessed via the URL: http://www.ars-grin.gov/duke.
20. Usia, T.; Iwataa, H.; Hiratsuka, A.; Watabe, T.; Kadota, S.; Tezuka, Y. *Phytomedicine* **2006**, *13*, 67.
21. Grosvenor, P. W.; Gothard, P. K.; McWilliam, N. C.; Supriono, A.; Gray, D. O. *J. Ethnopharmacol.* **1995**, *45*, 75.
22. Batubara, I.; Mitsunaga, T.; Ohashi, H. *J. Wood Sci.* **2009**, *55*, 230.
23. Chen, B.; Wild, D.; Guha, R. *J. Chem. Inf. Model.* **2009,** *49*, 2044–2055: accessed via URL: http://pubchem.ncbi.nlm.nih.gov.
24. Dictionary of Natural Products (DNP) 9.1, CRC Press, 2010; http://dnp.chemnetbase.com).
25. Ehrman, T. M.; Barlow, D. J.; Hylands, P. J. *J. Chem. Inf. Model* **2007**, *47*, 264.
26. Breiman, L.; Friedman, J. H.; Olshen, R. A.; Stone, C. J. *Classification and Regression Trees*; Chapman & Hall/CRC: Boca Raton, FL, 2000. Chapter 4, pp. 103–104.
27. Broto, P.; Moreau, G.; Vandycke, C. *Eur. J. Med. Chem.Chim. Ther.* **1984**, *19*, 66.
28. Moran, P. A. P. *Biometrika* **1950**, *37*, 17.
29. Geary, R. C. *The incorporated statistician* **1954**, *5*, 115.
30. Schuur, J. H.; Selzer, P.; Gasteiger, J. *J. Chem. Inf. Comput. Sci.* **1996**, *36*, 334.
31. Hemmer, M. C.; Steinhauer, V.; Gasteiger, J. *Vib. Spectrosc.* **1999**, *19*, 151.
32. Randic, M. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 373.
33. Randic, M.; Razinger, M. *J. Chem. Inf. Comput. Sci.* **1995**, *35*, 594.
34. Mylari, B. L.; Armento, S. J.; Beebe, D. A.; Conn, E. L.; Coutcher, J. B.; Dina, M. S.; O'Gorman, M. T.; Linhares, M. C.; Martin, W. H.; Oates, P. J.; Tess, D. A.; Withbroe, G. J.; Zembrowski, W. J. *J. Med. Chem.* **2003**, *46*, 2283.
35. Aida, K.; Tawata, M.; Shindo, H.; Onaya, T.; Sasaki, H.; Yamaguchi, T.; Chin, M.; Mitsuhashi, H. *Planta Med.* **1990**, *56*, 254.
36. Lee, Y. S.; Kim, S. H.; Jung, S. H.; Kim, J. K.; Pan, C.; Lim, S. S. *Biol. Pharm. Bull.* **2010**, *33*, 917.
37. Kubo, I.; Murai, Y.; Soediro, I.; Soetarno, S.; Sastrodihardjo, S. *J. Nat. Prod.* **1991**, *54*, 1115.
38. Akhtar, M. S.; Khan, M. A.; Malik, M. T. *Fitoterapia* **2002**, *73*, 623.
39. Tezuka, Y.; Kasimu, R.; Basnet, P.; Namba, T.; Kadota, S. *Chem. Pharm. Bull.* **1997**, *45*, 1306.