Indian Journal of Chemistry
Vol. 47B, June 2008, pp. 957-962

## Note

# Quantitative structure-activity relationship studies of a new class of selective aldose reductase inhibitors

R Sharma*, P K Sahoo & S C Chaturvedi

School of Pharmacy, Devi Ahilya Vishwavidyalaya, Takshashila Campus, Khandwa Road, Indore 452 017, M. P., India

Email: sunnitesh06@gmail.com

*Received 18 July 2007; accepted (revised) 12 March 2008*

Quantitative structure activity relationship studies on a series of [1,2,4]Triazino[4,3-*a*] benzimidazole acetic acid derivatives with selective aldose reductase inhibitor activity is made using a combination of various physicochemical descriptors. This is a new class of triazino benzimidazole derivatives having selective activity with minimum toxic effects. Several significant equations with good co-efficient of correlation (≥ 0.854) are obtained; the best model is selected using predictive ability of equations. The model shows positive contribution of logP towards biological activity i.e., high hydrophobic nature of the molecules might be increasing the selective aldose reductase inhibitor activity.

**Keywords**: Benzimidazol; aldose reductase activity; Quantitative Structure-Activity Relationship

The estimated prevalence of diabetes mellitus among adults worldwide was 4.0% in 1995 and is expected to double by 2025 (Ref. 1). In spite of insulin treatment most diabetic patients eventually experience long-term diabetic complications, such as retinopathy, neuropathy, cataract and angiopathy. Although there is still no definite pathogenic link between hyperglycemia and diabetic complications, several mechanisms seem to be involved in the toxic effects caused by excess glucose[2,3]. Among well-examined factors are the activation of protein kinase C[4,5], enhanced protein glycation with the formation of advanced glycated end products (AGEs)[6,7], rise of oxidative stress[8, 9], and activation of the polyol pathway[10].

The polyol pathway was first implicated in the etiology of secondary complications of diabetes[11]. Aldose reductase (AR) is the first enzyme of this pathway and is widely distributed in mammalian tissues[12]. In the presence of NADPH, the enzyme converts glucose to sorbitol, which is only slowly metabolized to fructose by sorbitol dehydrogenase, the other enzyme in the pathway, with concurrent reduction of $NAD^+$. The activation of the polyol pathway, which occurs during hyperglycemia, brings about various metabolic imbalances in tissues that undergo insulin-independent uptake of glucose. In the ocular lens, hyper osmotic swelling is caused by the accumulation of sorbitol. In other tissues, the depletion of the cofactor NADPH used in the pathway results in the deactivation of glutathione reductase and nitric oxide synthase, leading to an increased susceptibility to oxidative stress, vascular derangement and a decrease in nerve conduction velocity. It has been shown that the oxidation of sorbitol catalyzed by sorbitol dehydrogenase increases the ratio of NADH: $NAD^+$, resulting in an increased lactate: pyruvate ratio and pseudohypoxia[13].

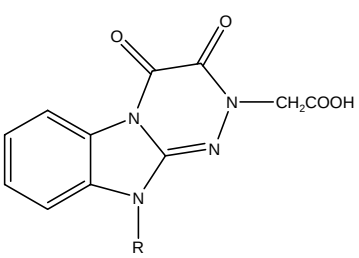
There exist a variety of structurally diverse aldose reductase inhibitors (ARIs). These compounds can be divided in to two general classes, those containing a carboxylic acid moiety and those having a cyclic imide represented by a spirohydantoin or related ring system[14-17]. Recently however, arylsulphonyl nitromethane has emerged as a new class. Although several ARIs have been tested in clinical trials on diabetic patients for more than 20 years, they still remain to be proven sufficiently effective[18]. Tolrestat, which was launched in 1989, was withdrawn in 1996, principally due to its low efficacy. Of the newer compounds, Zopolrestat and Zenarestat were withdrawn from clinical trials[19].

Aldose reductase (hALR2), a key component of the polyol pathway, has been a target for therapeutic intervention in the development of chronic diabetic complications. Although many potent aldose reductase inhibitors (ARIs) have been identified, the majority of these also inhibit aldehyde reductase (hALR1), a related enzyme involved in the detoxification of relative aldehydes. It is realized noteworthy to quantify aforementioned structure activity data and explore the nature of molecular interactions of these new class of ligands (**Table I**) with the aldose reductase enzyme to find out a novel, highly potent and selective aldose reductase inhibitor.

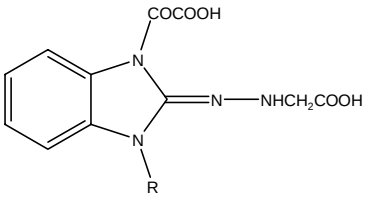
## Results

$pIC_{50}$ = 0.322(±0.199) LogP+0.000(±0.000)TotE1-0.189(±0.341)LUMO+3.927(±0.498)

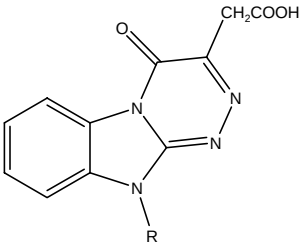
**Table I** — Structure and observed biological activity of series of [1,2,4]Triazino[4,3-*a*] benzimidazole acetic acid derivatives




**1 - 11**




**12 - 22**




**23 - 24**

| Compd | R | $IC_{50}$ [a] in μM | $pIC_{50}$ [b] |
|---|---|---|---|
| **1** | $CH_3$ | 24.8 | 4.6055 |
| **2** | $CH_2CH_2CH_3$ | 37.2 | 4.4294 |
| **3** | $CH_2C_6H_5$ | 0.36 | 6.4437 |
| **4** | $CH_2C_6H_4$-4-$CH_3$ | 13.3 | 4.8761 |
| **5** | $CH_2C_6H_4$-4-$OCH_3$ | 42.6 | 4.3705 |
| **6** | $CH_2C_6H_4$-4-Cl | 4.15 | 5.3819 |
| **7** | $CH_2C_6H_4$-4-F | 4.58 | 5.3391 |
| **8** | $CH_2C_6H_4$-4-$CF_3$ | 23.9 | 4.6216 |
| **9** | $CH_2C_6H_3$-3,4-$F_2$ | 4.42 | 5.3545 |
| **10** | $CH_2C_6H_3$-2-F-4-Br | 4.47 | 5.3496 |
| **11** | $CH_2COOH$ | 13.5 | 4.8696 |
| **12** | $CH_3$ | 108.6 | 3.9641 |
| **13** | $CH_2CH_2CH_3$ | 46.5 | 4.3325 |
| **14** | $CH_2C_6H_5$ | 4.50 | 5.3467 |
| **15** | $CH_2C_6H_4$-4-$CH_3$ | 45.9 | 4.3381 |
| **16** | $CH_2C_6H_4$-4-$OCH_3$ | 44.5 | 4.3516 |
| **17** | $CH_2C_6H_4$-4-Cl | 10.00 | 5.0000 |
| **18** | $CH_2C_6H_4$-4-F | 14.80 | 4.8297 |
| **19** | $CH_2C_6H_4$-4-$CF_3$ | 2.63 | 5.5800 |
| **20** | $CH_2C_6H_3$-3,4-$F_2$ | 9.72 | 5.0123 |
| **21** | $CH_2C_6H_3$-2-F-4-Br | 12.5 | 4.9030 |
| **22** | $CH_2COOH$ | 236 | 3.6270 |
| **23** | H | 35.9 | 4.4449 |
| **24** | $CH_3$ | 17.00 | 4.7695 |
| **25** | $CH_2C_6H_5$ | 5.44 | 5.2644 |

a: $IC_{50}$ (in μM) was the in vitro observed biological activity of compounds
b: Negative logarithmic value of $IC_{50}$

n=20, r=0.853864, $r^2$=0.729085, variance=0.10837, std=0.329196, F=14.353 … (Eqn.1)

Equation 1 fulfills many of the statistical validations such as the correlation coefficient; the cross validated squared correlation coefficient, standard deviation, bootstrapping squared correlation coefficient and chance. But the predictive residual sum of square standard error of prediction is less than 0.5 (0.35). The correlation accounted for more than 72.9% of the variance in the activity. The data showed an overall internal statistical significance level better than 99.9% as $F_{(3, 16\ \alpha\ 0.001)}$ = 14.353 which exceeds the

tabulated $F_{(3,\ 16\ \alpha\ 0.001)}$ = 9.01, the cross validated squared correlation coefficient ($Q^2$ = 0.639), the predictive residual sum of square $S_{PRESS}$ = 0.379), and the standard error of prediction ($S_{DEP}$ = 0.339) suggested good internal consistency as well as predictive activity of the biological activity with high logP.

$pIC_{50}$ = 0.278(±0.168) LogP-0.002 (±0.001) StrBE+4.179 (±0.361)

n=21, r=0.813953, $r^2$=0.662519, variance=0.141838, std=0.376615, F=17.6682 … (Eqn.2)

$pIC_{50}$=0.290(±0.168) LogP+0.000 (±0.000) TotE1 +4.178 (±0.364)

n=21, r=0.810943, $r^2$=0.657628, variance=0.143894, std=0.379334, F=17.2872 … (Eqn.3)

$pIC_{50}$ = 0.290 (±0.168) LogP+0.000 (±0.000) StrE+4.177 (±0.364)

n=21, r=0.810889, $r^2$=0.657541, variance=0.143931, std=0.379382, F=17.2805 … (Eqn.4)

$pIC_{50}$ = 0.285(±0.167)LogP-0.002(±0.001)StrBE-0.067(±0.106)DipL+4.546(±0.715)

n=21, r=0.825683, $r^2$=0.681752, variance=0.140703, std=0.375104, F=12.1392 … (Eqn.5)

Equations 2, 3, 4 and 5 were quite significant, which showed a bootstrapping squared correlation coefficient values such as 0.709, 0.729, 0.671 and 0.714 respectively. The inter correlation among the parameters of equation 2, 3, and 5 are 0.141, 0.105 and 0.105 respectively. But the intercorelation among the parameters of equation 4 are 0.142 for logP with Str.BE, 0.006 for logP with Dip.L and 0.383 for Str.BE with Dip.L In equations 2, 3, 4, and 5 one significant feature is that logP (thermodynamic descriptor) contribute positively towards biological activity where as other descriptors like Str.BE contribute negatively in equations 2 and 5, DipL also contribute negatively in equation 5. In equations 3 and 4 the contributions of Tot.E1 and StrE were very poor towards biological activity.

## Discussion

On the basis of various significant statistical validation data (**Table II**) equation 2 was selected as best model, which may be a true representation in order to explore the factors responsible for the selective aldose reductase inhibitor activity of the series of analogs. This may be helpful in designing more potent substituted [1,2,4]Triazino[4,3-*a*]benzimidazole acetic acid derivatives.

### Validation of the model

Equation 2 shows a better correlation coefficient (r=0.8139), which accounts for more than 66.25% of the variance in the activity, also the intercorrelation among the parameters is less (0.141). The equation shows that in the multi-variant model, the dependent variable can be predicted from a linear combination of the independent variables. The P-value is less than 0.001 for each physiochemical parameter involved in model generation. The data showed an over all internal statistical significance level better than 99.9% as it exceeded the tabulated $F_{(3,\ 17\ \alpha\ 0.001)}$ = 8.73. The model was further tested for the outlier by the Z-score method and no compound was found to be an outlier), which suggested that the model is able to explain the structurally diverse analog and is helpful in designing more potent compounds using physiochemical parameters. The leave-one-out cross validation method was employed for the prediction of activity (**Figure 1**). The cross-validated squared correlation coefficient (in the biological activity data of leave-one-out) ($Q^2$ = 0.534), predictive residual sum of square ($S_{PRESS}$ = 0.442), and standard error of

**Table II** — QSAR statistics of significant equations

| Model | N Train. | N Test | NV | r | $r^2$ | $Q^2$ | Chance | $r^2$-pred. | Spress | SDEP |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 20 | 5 | 3 | 0.8538 | 0.7290 | 0.63 | 0.01 | 0.35 | 0.379 | 0.33 |
| 2 | 21 | 4 | 2 | 0.8139 | 0.6625 | 0.53 | 0.01 | 0.50 | 0.442 | 0.40 |
| 3 | 21 | 4 | 2 | 0.8109 | 0.6576 | 0.37 | 0.01 | 0.63 | 0.513 | 0.47 |
| 4 | 21 | 4 | 2 | 0.8108 | 0.6575 | 0.37 | 0.01 | 0.63 | 0.514 | 0.47 |
| 5 | 21 | 4 | 3 | 0.8256 | 0.6817 | 0.31 | 0.01 | 0.54 | 0.550 | 0.49 |

N Train: Number of compounds in training set, N Test: Number of compounds in test set, NV: Number of independent variables, r: Coefficient of correlation, $r^2$: Coefficient of determination $Q^2$: cross-validated squared correlation coefficient, $r^2$ pred.: Predicted coefficient of correlation, Spress: Predictive residual sum of square, SDEP: Standard error of prediction

prediction ($S_{DEP}$ = 0.409) suggested a good internal consistency as well as predictive ability of the biological activity with low $S_{DEP}$. The $r^2_{bs}$ is at par with the conventional squared correlation coefficient ($r^2$). Randomized biological activity results were not based on the correlation. The robustness and wide applicability of the model were further explained by significant $r^2_{pred}$ value (0.50) (**Figure 2** and **Table II**). In general, the model fulfils the statistical validation criteria to a significant extent to be a useful theoretical base for proposing more active compound. In equation 2 logP contributed positively where as Str.BE contributed negatively towards biological activity. LogP is representative of atoms of hydrophobic nature in the molecules and suggests that substitution of groups, which are high hydrophobic in nature, might increase the biological activity. Thus, reimproving the logP characteristics of the molecule increases the selective aldose reductase inhibitor activity. Whereas minimizing the property like Str.BE which is helpful for rationalizing the interaction between molecule and receptor surface. The study revealed that distal end substitutions might interact with a hydrophobic pocket at receptor site, hence increasing hydrophobicity of the substituent increase the binding capacity between molecule and receptor surface which potentiate the selectivity as well as activity.

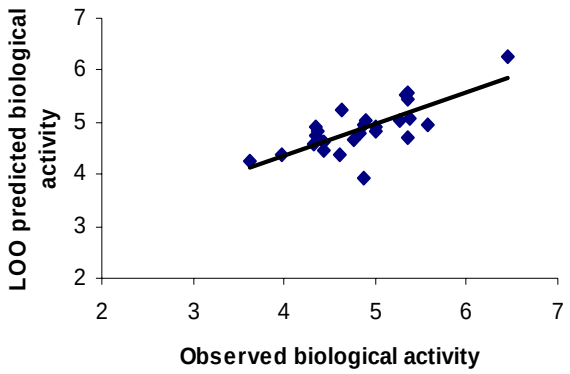



**Figure 1** — Plot between LOO predicted $pIC_{50}$ and observed $pIC_{50}$ values of compounds of training set for equation 2

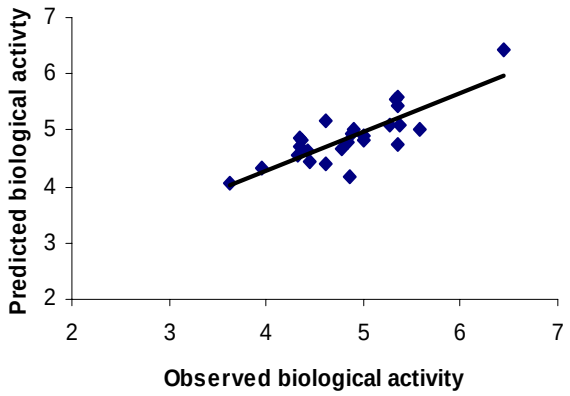



**Figure 2** — Plot between predicted $pIC_{50}$ and observed $pIC_{50}$ values of compounds of training set for equation 2

## Materials and methods

### Data set

The aldose reductase inhibitor activity data of [1,2,4]Triazino[4,3-*a*]benzimidazole acetic acid derivatives were taken from the reported work of Settimo *et al*[20]. The biological activity data ($IC_{50}$ in μm) were converted to negative logarithmic dose ($pIC_{50}$) for quantitative structure activity analysis.

### Geometry optimization

The molecular structures of all 25 compounds were sketched using the Chemdraw Ultra (Version 8.0) software and energy minimized via MOPAC with energy tolerance value of root mean square gradient 0.001 kcal/mol and maximum number of iteration set to 1000. Conformational search of each energy-minimized structure was performed using the stochastic approach which is similar to the RIPS method. All conformers generated for each structure were analyzed in conformational geometrics panels with great care, and the lowest energy conformation of each structure was selected and added to a molecular database to compute various physico-chemical properties. The descriptor values used in the model generation are shown in the **Table III**.

### Statistical methods and molecular descriptors

The series was divided in to a training set of 21 compounds and a test set of four compounds carried out automatically by the VALSTAT software (**Tables IV and V**). The sequential multiple linear regression analysis method was employed. In sequential multiple linear regression, the program searches for all permutations and combinations sequentially for the data set. In this case it searched for 1,25,600 combinations and gave 10 multi-variant equations based on squared correlation coefficient. The ± data with in the parentheses are the standard deviations associated with the coefficient of descriptors in regression equations. The best model

**Table III** — Calculated values of independent variables

| Compd | LogP | Str. E | Str.BE | Tot. E | Dip.L | LUMO E |
|---|---|---|---|---|---|---|
| **1** | 0.7898 | 7.62887 | -1.7791 | 38.0291 | 4.67541 | -0.6250 |
| **2** | 1.6140 | 8.18515 | -1.8978 | 41.2021 | 7.38977 | -0.5308 |
| **3** | 2.2748 | 2592000 | -716.6 | 2594470 | 3.12539 | -0.4494 |
| **4** | 2.7619 | 7.92579 | -1.7423 | 35.2303 | 5.37921 | -0.6341 |
| **5** | 2.1484 | 388.894 | -22.288 | 768.095 | 6.71199 | -1.1562 |
| **6** | 2.8330 | 36748.5 | -53.603 | 37289.7 | 6.89207 | -2.0544 |
| **7** | 2.4329 | 860074 | -306.17 | 860817 | 2.24122 | -0.5139 |
| **8** | 3.1959 | 23684.1 | -37.583 | 23989.3 | 6.70292 | -0.7406 |
| **9** | 2.5910 | 445732 | -237.75 | 448276 | 11.4074 | -2.0125 |
| **10** | 3.2618 | 68410.5 | -218.79 | 71352.7 | 3.50311 | -0.8639 |
| **11** | -0.1916 | 17371.8 | -15.162 | 17712.7 | 2.43480 | -1.1303 |
| **12** | 0.5351 | 5.80272 | -1.8593 | 68.9341 | 7.28599 | -0.4083 |
| **13** | 1.3593 | 5.3207 | -1.4613 | 49.1808 | 7.19586 | -0.4495 |
| **14** | 2.0201 | 5.49892 | -1.7844 | 38.2714 | 5.05971 | -0.4446 |
| **15** | 2.5072 | 5.24168 | -1.354 | 38.0078 | 5.01210 | -0.3633 |
| **16** | 1.8937 | 6.65041 | -1.6197 | 68.0748 | 6.71447 | -0.6010 |
| **17** | 2.5783 | 5.59822 | -1.6605 | 43.7304 | 3.13483 | -0.5134 |
| **18** | 2.1782 | 5.85121 | -1.4541 | 37.8953 | 6.15278 | -0.4559 |
| **19** | 2.9412 | 8.1529 | -1.4131 | 48.0357 | 5.38989 | -0.7689 |
| **20** | 2.3363 | 7.45931 | -1.8773 | 47.6272 | 3.47308 | -0.7301 |
| **21** | 3.0071 | 7.09369 | -2.0858 | 39.7411 | 5.11954 | -0.7437 |
| **22** | -0.4463 | 5.81768 | -1.4822 | 61.932 | 5.17877 | -0.7395 |
| **23** | 0.9445 | 9.39299 | -1.3358 | 34.561 | 8.62682 | -1.2655 |
| **24** | 1.7328 | 8.22259 | -1.4649 | 39.9458 | 6.50269 | -1.0621 |
| **25** | 3.2178 | 8.15495 | -1.3927 | 37.9949 | 8.81481 | -1.1011 |

Str.E: Sterching Energy, Str.BE: Stretching bending energy, Tot. E: Total energy,
Dip. L dipole length, LUMO E: Lowest unoccupied molecular orbital energy

**Table V** — Predicted biological activity and LOO predicted activity with their variance in comparison to the observed biological activity of equation 2 (Test Set)

| Comp. | Obs. Activity | Pred. Activity | Variance | LOO pred. | Variance |
|---|---|---|---|---|---|
| **5** | 4.3705 | 4.8261 | -0.4556 | 4.8261 | -0.4556 |
| **7** | 5.3391 | 5.5447 | -0.2056 | 5.5447 | -0.2056 |
| **10** | 5.3496 | 5.5783 | -0.2287 | 5.5783 | -0.2287 |
| **17** | 5.000 | 4.8991 | 0.1009 | 4.8991 | 0.1009 |

Obs. Activity: Observed biological activity, Pred. Activity: Predicted biological activity, LOO pred.: Leave one out predicted biological activity

**Table VI** — Inter correlation matrix of the independent descriptor for equation 2

| | LogP | Str. BE |
|---|---|---|
| LogP | 1.000000 | |
| Str.BE | 0.141725 | 1.000000 |

was selected from the various statistically significant equations on the basis of the observed squared correlation coefficient ($r^2$), variance (v), standard deviation (std.) the sequential Fischer test (F), the Bootstrapping $r^2$, chance, $Q^2$ value, $S_{press}$ value, standard deviation of error prediction (SDEP) and the predictive squared correlation coefficient of the test set ($r^2$ pred.)[21], (**Table VI**).

## Conclusion

High lipophilic nature of the molecule and low electrostatic potential energy are favourable for the selective aldose reductase inhibitor activity. Thus, modification in structure to improve lipophilic character and electrostatic potential energy might result in a more potent selective aldose reductase inhibitor.

## Acknowledgement

The authors are grateful to the Head, School of Pharmacy, Devi Ahilya Vishwavidyalaya, Indore, Madhya Pradesh, India for providing facilities for this work.

## References

1 King H, Aurebt R E & Herman W H, *Diabetes Care*, 21, **1998**, 1414.

2 Porte D & Schwartz M W, *Science*, 272, **1996**, 699.

3 Sima A A F & Sugimoto K, *Diabetologia*, 50, **1999**, 21.

4 Koya D & King G L, *Diabetes*, 47, **1998**, 859.

5 Ishii H, Jirousek M R, Koya D, Takagi C, Xia P, Clermont A, Bursell S E, Kern T S, Ballas L M, Heath W F, Stramm L E, Feener E P & King G L, *Science*, 272, **1996**, 728.

6 Friedman E A, *Diabetes Care*, 22(2), **1999**, B65.

7 Brownlee M, *Metabolism*, 49(1), **2000**, 9.

8 Baynes J W & Thorpe S R, *Diabetes*, 48, **1999**, 1.

9 Betteridge D J, *Metabolism*, 49(1), **2000**, 3.

10 Yable-Nishimura C, *Pharmacol, Rev*. 50, **1998**, 21.

11 Tomlinson D R, Stevens E J & Diemel L T, *Trends Pharmacol Sci*, 15, **1994**, 293.

12 Carper D, Wistow G, Nishimura C N, Graham C, Watanabe K, Fuji Y, Hayaishi H & Hayaishi O, *Exp Eye Res*, 49, **1998**, 377.

13 Williamson J R, Chang K, Frangos M, Hasan K S, Ido Y, Kawamura T, Nyengaard J R, Van-den-Enden M, Kilo C & Tilton R G, *Diabetes*, 42, **1993**, 801.

14 Wrobel J, Millen J, Sredy J, Dietrich A, Gorham B J, Malamas M, Kelly J M, Bauman J G, Harrison M C, Jones L R, Guinosso C & Sestanj K, *J Med Chem*, 34, **1991**, 2504.

15 Yamagishi M, Yamada Y, Ozaki K, Asao M, Shimizu R, Suzuki M, Matsumoto M, Matsumoto Y & Matsumoto K J, *J Med Chem*, 35, **1992**, 2085.

16 Costantino L, Rastelli G, Vianello P, Cignarella G & Barlocco D, *Med Res Rev*, 19, **1999**, 3.

17 Miyamoto S, *Expert Opin Ther Patents*, 12, **2001**, 621.

18 Preifer M A, Schumer M P & Gelber D A, *Diabetes*, 46(2), **1997**, 582.

19 Mylari B L, Larson E R, Beyer T A, Zembrowski W J, Aldinger C E, Dee M F, Siegel T W & Singleton D H, *J Med Chem*, 34, **1991**, 108.

20 Settimo F D, Primofiore G, Settimo A D, Motta C L, Taliani S, Simorini F, Novellino E, Greco G, Lavecchia A & Boldrini E, *J Med Chem*, 44, **2001**, 4359.

21 Hanumantharao P, Sambasivarao S V, Soni L K, Gupta A K & Kaskhedikar S G, *Bioorg Med Chem Lett*, 15, **2005**, 3167.