

Quantitative structure–activity relationship of spirosuccinimide type aldose reductase inhibitors diminishing sorbitol accumulation in vivo

Kwangseok Ko,^a Hoshik Won^{b,*} and Youngdo Won^{a,*}

^aDepartment of Chemistry, Hanyang University, Seoul 133-791, Republic of Korea

^bDepartment of Applied Chemistry, Hanyang University, Ansan 426-791, Republic of Korea

Received 15 November 2005; revised 14 December 2005; accepted 15 December 2005

Available online 18 January 2006

Abstract—Racemate physicochemical descriptors are employed to probe the quantitative structure–activity relationship of spirosuccinimide type aldose reductase inhibitors and the in vivo inhibitory activity of sorbitol accumulation. The in vivo activity data include the percent inhibition and ED₅₀ assay results on the literature. The derived QSAR equations show that the hydrophobic character of aldose reductase inhibitor is the major contributing factor to enhance in vivo activity. As the hydrophobicity of compounds is related to both the membrane permeability and the binding affinity to the aldose reductase, its contribution to the pharmacokinetic behavior is further scrutinized by evaluating p*K*_a and the Caco-2 cell permeability. The high correlation between ED₅₀ and the Caco-2 cell permeability of in vitro active compounds indicates that the membrane permeability is essential for in vivo efficacy.

© 2005 Elsevier Ltd. All rights reserved.

1. Introduction

The hyperglycemia observed in diabetes mellitus seems to be the primary instigator for the pathogenesis of long-term diabetic complications such as retinopathy, neuropathy, nephropathy, and cataracts.^{1,2} The elevated glucose concentration in blood activates the polyol pathway, of which the first enzyme is aldose reductase (AR) reducing glucose into sorbitol.^{3,4} As the excessive accumulation of intracellular sorbitol through the polyol pathway is linked to the pathogenesis of diabetic complications, prevention of sorbitol accumulation by inhibiting the AR activity would be an effective treatment.^{5,6}

Although a large number of AR inhibitors have been discovered to show significant activity in vitro and in

animal models,^{7–17} most of them have been withdrawn from clinical trial due to lack of efficacy or adverse side effects.^{18–20} The promising compound in the in vitro assay often failed in the in vivo test because of pharmacokinetic problems such as absorption, distribution, metabolism, and excretion (ADME). For instance, tolrestat and ponalrestat reached the stage of clinical trial but they were withdrawn due to low efficacy.^{19,20} Recently, fidarestat showed encouraging results normalizing the erythrocytic sorbitol contents in neuropathy patients with no significant side effects.²¹ However, it is still not recognized as an effective therapeutic agent for worldwide use.

In search for a more potent and effective AR inhibitor, Negoro et al. developed a series of spirosuccinimide-fused tetrahydropyrrolo[1,2-*a*]pyrazine derivatives.²² We analyzed the quantitative structure–activity relationship (QSAR) of the tetrahydropyrrolo[1,2-*a*]pyrazine derivatives against their in vitro activity data.^{23,24} The spirosuccinimide type AR inhibitors were prepared as racemates rather than optically pure compounds, except a few cases. In order to examine the in vitro assay result of racemates, we devised QSAR descriptors suitable for racemic mixtures. The racemic descriptor, so called the *RS*-descriptor, was

Keywords: QSAR; Aldose reductase inhibitor; Sorbitol accumulation; Caco-2 cell permeability.

* Corresponding authors. Tel.: +82 31 400 5497 (H.W.); tel.: +82 2 2220 0944 (Y.W.); e-mail addresses: hswon@hanyang.ac.kr; won@hanyang.ac.kr

derived as the simple arithmetic mean of the *R*-descriptor (for the (*R*)-enantiomer) and the *S*-descriptor (for the (*S*)-enantiomer). The QSAR model derived with the *RS*-descriptors was superior to those obtained with either the original *R*-descriptors or the *S*-descriptors.

Here we employ the *RS*-descriptors to investigate QSAR of spirosuccinimide-fused tetrahydropyrrolo[1,2-*a*]pyrazine-1,3-dione derivatives preventing sorbitol accumulation in vivo. The QSAR model would identify the physicochemical property of candidate compounds enhancing the pharmacokinetic behavior and the in vivo potency. We also evaluate pK_a and the cell permeability of the inhibitor compounds to elucidate the pharmacokinetic behavior.

2. Results and discussion

The in vitro and in vivo activity data of the series of spirosuccinimide-fused tetrahydropyrrolo[1,2-*a*]pyrazine-1,3-dione derivatives are obtained from the assay results of Negoro et al.,²² and collected in Table 1. IC_{50} refers to the concentration of the compound required for 50% inhibition of the porcine lens AR activity. ED_{50} is the dose causing 50% decrease of the sorbitol accumulation in the sciatic nerve of diabetic rats in vivo. The percent inhibition of sorbitol accumulation was measured at the dose of 30 mg/kg/day and the test compound was orally given once a day for 5 days.²² The percent inhibition value of over 100% means that the test compound decreases the sorbitol content in the sciatic nerve of diabetic rats below that in the control, the sciatic nerve of nondiabetic rats. The percent inhibition value and the negative logarithm of ED_{50} (pED_{50}) are the dependent variables in our QSAR analyses.

We first built the QSAR equations for the percent inhibition of sorbitol accumulation in the sciatic nerve of diabetic rats (Table 1). Following the same protocol employed to derive the QSAR equations for the in vitro pIC_{50} data,^{23,24} we determined the optimal number of terms of the in vivo QSAR equation based on the statistics of resultant QSAR models. We sequentially applied the genetic function approximation (GFA) technique²⁵ with the *RS*-descriptor set by limiting the number of terms from one through six and generated six QSAR models. Among those best QSAR equations, the five term QSAR equation shows the lowest *LOF* score (0.187) and the largest *F* value (51.822). The extra descriptor to the five term equation slightly improves the statistics; the regression coefficient of the six term equation is 0.975 and the standard deviation is 0.261. However, the six term equation has a higher *LOF* score (0.225) and a lower *F* value (48.875). Furthermore, the sequential *F* value (F_s) for the sixth term addition is 2.954. Therefore, we fixed the number of descriptors to five in the QSAR equations for the percent inhibition activity. The best five term QSAR equation is shown in Eq. 1 with the statistics.

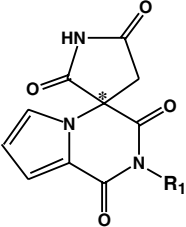
$$\begin{aligned} \% \text{ Inhibition} = & 0.718 \text{ PNSA}_1 - 0.528 V_{\text{NCOS}}^2 - 0.342 \mu_z \\ & - 0.262 F_{\text{H}_2\text{O}}^2 + 0.254 S_z \\ n = 22, r = 0.970, q^2 = 0.911, LOF = 0.187, s = 0.276, \\ F = 51.822(2.852), F_s = 9.881 \end{aligned} \quad (1)$$

where *n* is the number of compounds, *r* is the regression coefficient, q^2 is the cross-validated r^2 from the leave-one-out procedure, *LOF* is Friedman's 'lack of fit' score, and where *s* is the standard deviation of estimation. *F* is the *F* value, a measure of statistical significance level of the model, and the *F* value given in parentheses is that of 95% significance level. F_s is the sequential *F* value, which is used to compare two regression models with different number of descriptors.²⁴

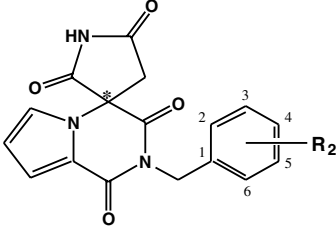
We also examined the performance of *RS*-descriptors in the QSAR analyses of the in vivo activity data. We separately derived five term QSAR equations with *R*-descriptors and *S*-descriptors, and called them *R*-equation and *S*-equation, respectively (data not shown). The regression coefficient is 0.952 and the cross-validated r^2 is 0.823 for the *R*-equation. Those statistical values are 0.970 and 0.901, respectively, for the *S*-equation. The racemic QSAR model derived with the *RS*-descriptors shows a higher regression coefficient and predictive power than the original *R*-equation and the *S*-equation. The *RS*-descriptors are suitable to the QSAR analysis of the in vivo activity data as well as the in vitro data obtained with racemates.

The first entry of the five term QSAR equation is the partial negative surface area, PNSA_1 . PNSA_1 is defined as the sum of the solvent-accessible surface areas of all negatively charged atoms.²⁶ PNSA_1 and the percent inhibition most strongly correlate with the correlation coefficient of 0.838. $A \log P$ is the direct measure of the hydrophobicity of the compound.²⁷ The strong correlation between $A \log P$ and the percent inhibition (the correlation coefficient of 0.658) indicates that $A \log P$ would also positively contribute to the percent inhibition. However, $A \log P$ is filtered out in the descriptor screening process as it shows strong correlation with PNSA_1 . The correlation coefficient is 0.746. The positive contribution of PNSA_1 and $A \log P$ suggests that the hydrophobic character of the compound is very important for high in vivo activity.

Previously, we showed that the in vitro AR inhibitory activity (pIC_{50}) of 30 tetrahydropyrrolo[1,2-*a*]pyrazine-1,3-dione derivatives was primarily governed by the hydrophobicity of compounds.²⁴ The in vitro activity and the in vivo percent inhibition activity show strong correlation with the correlation coefficient of 0.789. The hydrophobicity of compounds plays an essential role in the direct hydrophobic interaction with AR in vivo as well as in vitro. However, the hydrophobic character is related not only to the direct hydrophobic interaction with AR but also to the membrane permeability of the compound. The pharmacokinetic behavior should not be ruled out in the QSAR model of in vivo activity.

Table 1. Predicted and observed activity data of 2-substituted-1,2,3,4-tetrahydropyrrolo[1,2-*a*]pyrazine-4-spiro-3'-pyrrolidine-1,2',3,5'- tetrones^a


Compound	R ₁	R ₂	pIC ₅₀ ^b	% Inhibition ^c		pED ₅₀ ^c	
				Observed	Predicted ^d	Observed	Predicted ^f
1	H	—	0.201	33.2	37.9		
2	CH ₃ -	—	0.553	51.2	50.2		
3	C ₆ H ₅ -	—	0.699	27.2	25.3		
4	C ₆ H ₅ -(CH ₂) ₂ -	—	0.569	15.3	15.0		



5	—	H	1.004	69.1	63.4		
6	—	2-F	1.215	98.0	91.8	5.419	5.065
7	—	4-F	0.921	58.8	71.4		
8	—	2-Cl	1.432	97.0	83.8	4.790	5.028
9	—	3-Cl	1.456	99.4	100.7	4.974	5.147
10	—	4-Cl	1.420	104.6	88.1	5.378	5.252
11	—	2-Br	1.284	82.6	84.2		
12	—	3-Br	1.432	102.1	96.8		
13	—	4-Br	1.328	108.1	99.6	5.923	5.587
14	—	4-CH ₃	1.284	65.8	69.3		
15	—	4-OCH ₃	1.337	46.7	62.3		
16	—	4-CF ₃	1.097	88.4	91.5		
17	—	4-NO ₂	0.959	102.2	108.6	4.533	4.433
18	—	4-NH ₂	0.854	33.4	31.1		
19	—	2,4-F ₂	1.143	98.0	95.3	4.823	5.270
20	—	3,4-Cl ₂	1.638	111.8	115.6	5.815	5.819
21	—	2-F, 4-Cl	1.387	109.5	112.6	5.706	5.763
22	—	2-F, 4-Br	1.347	114.6	122.4	6.146	6.143

^a The observed in vitro and in vivo activity values are obtained from Ref. 22.^b pIC₅₀ = -log(IC₅₀), where the IC₅₀ value is in μM.^c Test compounds were orally given once a day for 5 days at 30 mg/kg/day.^d % Inhibition is predicted by the five term equation (Eq. 1).^e pED₅₀ = -log(ED₅₀), where the ED₅₀ value is in mol/kg/day.^f pED₅₀ is predicted by the two term equation (Eq. 2).

The second entry is the noncommon overlap steric volume, V_{NCOS} . The negative coefficient of V_{NCOS}^2 implies that the molecular shape should match that of the reference molecule (compound 22). The molecular shape descriptor would reflect the binding conformation of the compound. In the previous QSAR analysis of the in vitro activity, V_{NCOS} is one of the most important contributing descriptors to the inhibitory activity against AR. The V_{NCOS}^2 term indicates that both in vitro and in vivo activities require the binding affinity of the compound to AR.

The z-component dipole moment, μ_z , negatively contributes to the percent inhibition. It indicates that a com-

pound with the larger dipole moment in the negative z-direction is expected to have greater activity. The in vitro QSAR analysis shows that the contribution of hydrophobic character to the inhibitory activity is modulated by the polar descriptors such as the difference in charged partial surface area (DPSA₁) and the x-component dipole moment (μ_x). The correlation coefficient between μ_z and μ_x (the negative contributor to the in vitro activity) is 0.556. μ_z reflects the polar modulation on the hydrophobic character of the compound binding into the AR active site.

$F_{\text{H}_2\text{O}}$ is the aqueous desolvation free energy of the molecule. The solubility of the compound moderately con-

tributes to the in vivo activity as indicated by the correlation coefficient -0.296 between $F_{\text{H}_2\text{O}}^2$ and the percent inhibition. The square of the aqueous desolvation free energy negatively contributes to in vivo inhibition. The lower desolvation free energy for the higher activity implies that in vivo potency requires both the binding affinity to AR and the membrane permeability.

The last entry of the five term QSAR equation is S_z , the length of the molecule in the z-direction.²⁸ S_z positively contributes to the percent inhibition. The positive contribution of S_z implies that the extended conformation is preferred for the biological activity.

The correlation matrix of the selected descriptors in the QSAR (Eq. 1) is given in Table 2. The maximum value of the intercorrelation coefficient is 0.369 between V_{NCOS}^2 and S_z . All the intercorrelation coefficients of the descriptors are less than 0.4, which indicates that the five descriptors selected construct reasonably independent space for the QSAR model. The percent inhibition activities predicted by Eq. 1 are listed in Table 1 and are compared with the observed activities in Figure 1. The standard deviation is 0.276 and the cross-validated r^2 is 0.911.

We also investigated the in vivo activity of spirosuccinimide-fused tetrahydropyrrolo[1,2-*a*]pyrazine-1,3-dione derivatives by using pED_{50} as the dependent variable for the QSAR analysis. The same GFA protocol was used to generate a QSAR equation set in the reasonably independent *RS*-descriptor space. The best QSAR equation is shown in Eq. 2 with the statistical measures.

$$\text{pED}_{50} = 0.780 \text{ Alog } P - 0.409 V_{\text{NCOS}}^2$$

$$n = 10, r = 0.892, q^2 = 0.598, \text{LOF} = 0.510, s = 0.512,$$

$$F = 13.647(4.737), F_s = 5.715. \quad (2)$$

We obtained one through three term QSAR equations by limiting the number of terms in the GFA procedure. We select the two term equation as the best QSAR model for further analyses because it has the lowest *LOF* of 0.510 and the largest *F* value of 13.647. Although the third term addition improves the regression coefficient to 0.926, the three term equation fails in the sequential *F* test. The F_s value for the third term S_y^2 addition is 2.602, which is lower than the 95% significance limit (4.757). The *LOF* score of the three term equation is 0.801 and the *F* value is 12.047.

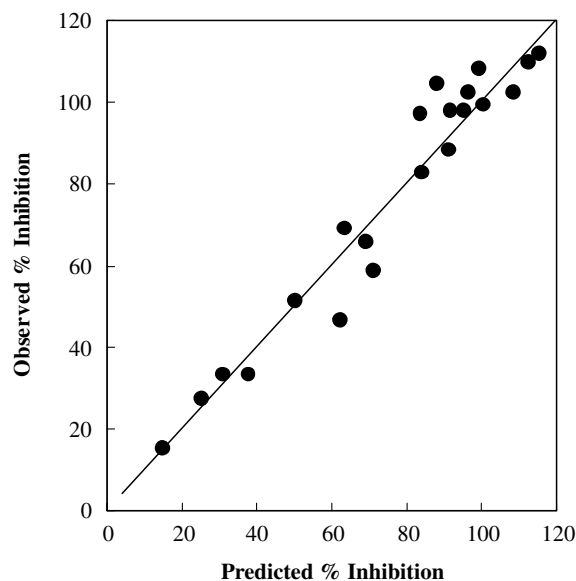


Figure 1. The percent inhibition values of tetrahydropyrrolo[1,2-*a*]pyrazine-1,3-dione derivatives predicted by the five term QSAR equation are compared to the assay data.

Eq. 2 shows that $\text{Alog } P$ is the primary term and V_{NCOS}^2 negatively contributes to the in vivo activity. As in the QSAR model of the percent inhibition, the hydrophobicity is the key factor in QSAR of pED_{50} . The QSAR model also indicates that the molecular shape of the reference compound **22** is well optimized and deviation from the reference structure is unfavorable for the inhibitory activity. $\text{Alog } P$ and V_{NCOS}^2 are well independent as indicated by the intercorrelation coefficient -0.032 . The correlation coefficients of $\text{Alog } P$ and V_{NCOS}^2 with pED_{50} are 0.793 and -0.434 , respectively, which are very close to the coefficients of Eq. 2. The pED_{50} values predicted by the two term QSAR equation are listed in Table 1 and are plotted against the observed activities in Figure 2.

$\text{Alog } P$, PNSA_1 , and most surface area descriptors correlate strongly with each other and reflect the hydrophobicity of the molecule. The hydrophobic descriptors primarily contribute to the AR inhibitory activities. It is likely that the hydrophobic character of compounds has the dual role in both the pharmacokinetic behavior and the binding affinity to AR. While the contributing factor for in vitro IC_{50} is the indicator of the binding affinity to AR, those for in vivo ED_{50} and the percent inhibition play an extra role in the pharmacokinetic behavior like absorption, membrane transport, and distribution in the biological system.

Table 2. Correlation matrix for the terms of the QSAR equation for the percent inhibition

	PNSA_1	V_{NCOS}^2	μ_z	$F_{\text{H}_2\text{O}}^2$	S_z	% Inhibition
PNSA_1	1.000					
V_{NCOS}^2	-0.172	1.000				
μ_z	0.273	-0.249	1.000			
$F_{\text{H}_2\text{O}}^2$	-0.135	0.027	-0.211	1.000		
S_z	0.347	0.369	0.310	0.018	1.000	
% Inhibition	0.838	-0.480	0.120	-0.296	0.197	1.000

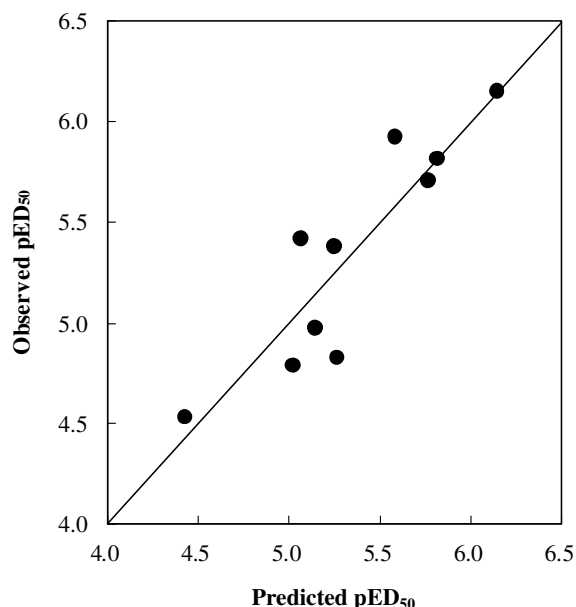


Figure 2. pED₅₀ of tetrahydropyrrolo[1,2-*a*]pyrazine-1,3-dione derivatives predicted by the two term QSAR equation are compared to the assay data.

Although ED₅₀ is assayed for only active compounds in the in vitro AR inhibitory assay (pIC₅₀ ≥ 1), there is substantial variation in the in vivo activity of the 10 training compounds. The correlation coefficient between pED₅₀ and pIC₅₀ is only 0.503. It suggests that the in vitro affinity for AR is not sufficient but necessary for the in vivo activity. The hydrophobic descriptors of Eq. 1 and 2 must implicate the ADME profile of the compound as well as the AR binding affinity.

We further considered pK_a and cell permeability as factors establishing the ADME properties of compounds. The cell permeability is the direct measure for absorption of the compound. We utilized the pK_a DB module of the ACD/MedChem Advisor program²⁹ to evaluate the pK_a value of the ionizable moiety of the compounds. The ACD/pK_a DB program calculates accurate acid–base ionization constants at 25 °C and zero ionic strength in aqueous solutions.²⁹ We also predicted the Caco-2 cell permeability of each compound by using the PreADME program, which was implemented with the back-propagation method on the artificial neural network.³⁰

pK_a value and the Caco-2 cell permeability are evaluated for 10 tetrahydropyrrolo[1,2-*a*]pyrazine-1,3-dione derivatives whose pED₅₀ was determined and two acetic acid derivatives, ponalrestat and tolrestat. The pK_a values and the Caco-2 cell permeabilities are listed in Table 3 along with in vitro and in vivo activities.

While ponalrestat and tolrestat show the similar in vitro activities as the active spirosuccinimide AR inhibitors, they have much lower in vivo activities (Table 3). Ponalrestat and tolrestat have pK_a values of 3.360 and 3.486, respectively, due to the carboxylic moiety and are

expected to be ionized at the physiological pH. The acidic group may act adversely on the membrane transport to result in poor in vivo activity. On the other hand, the 1' nitrogen in the succinimide ring of spirosuccinimide-fused tetrahydropyrrolo[1,2-*a*]pyrazine-1,3-dione derivatives yields a pK_a of around 10.5 and it maintains the neutral form at the physiological pH.

The predicted Caco-2 cell permeability, the last column of Table 3, clearly demonstrates that the absorption characteristic strongly correlates with the in vivo activity. While the Caco-2 cell permeability of ponalrestat and tolrestat is 0.464 and 1.343, respectively, it is greater than 3.0 for the spirosuccinimide derivatives, except for compound 17. The Caco-2 cell permeability of compound 17 is 0.210 due to the polar *p*-NO₂ group.

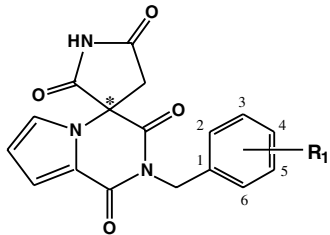
Figure 3 graphically shows the dependence of pED₅₀ on the Caco-2 cell permeability. The correlation coefficient of pED₅₀ and the Caco-2 cell permeability is 0.638 including the outlier compound 8. Excluding the outlier, it is 0.863. The in vivo activity of candidate compounds depends on both the affinity to the AR active site and the pharmacokinetic property. Although we included only 10 active compounds in the QSAR analysis, we observed high correlation between pED₅₀ and the Caco-2 cell permeability. The polarization states as well as the hydrophobicity of AR inhibitors are very important for the potent agent to prevent sorbitol accumulation in vivo.

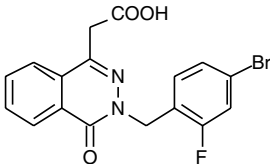
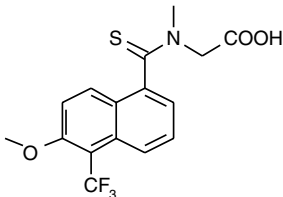
3. Conclusion

The quantitative structure–activity relationship of spirosuccinimide type AR inhibitors was investigated for two sets of in vivo inhibitory activities, the percent inhibition with 22 compounds and ED₅₀ with 10 compounds. As the compounds were assayed as the racemate, it is advantageous to employ physicochemical descriptors incorporating the contribution of each enantiomer. The racemic *RS*-descriptor, the simple arithmetic mean of *R*-descriptor and *S*-descriptor, turned out to be suitable for the in vivo QSAR as it did for the in vitro QSAR.

The resultant in vivo QSAR equations show that the hydrophobicity is the major factor for inhibition of sorbitol accumulation and that compound 22 is well optimized for the AR inhibitory activity. The same holds true for the in vitro IC₅₀ QSAR analysis performed previously. The hydrophobic character is desired for both absorption at the cell membrane and binding affinity at the AR active site. The QSAR models by themselves do not discriminate the hydrophobic interaction for AR binding from that for pharmacokinetic ADME properties.

We also evaluated pK_a and the Caco-2 cell permeability for 10 active spirosuccinimide-fused tetrahydropyrrolo[1,2-*a*]pyrazine-1,3-dione derivatives and compared them to those of two well-known AR inhibitors, ponalrestat and tolrestat. Low in vivo efficacy of ponalrestat and tolrestat is due to low pK_a and the low Caco-2 cell permeability of the compounds. Spirosuccinimide compounds have pK_a well

Table 3. pK_a and the Caco-2 cell permeability evaluated for AR inhibitors


Compound	R ₁	pIC ₅₀ ^a	pED ₅₀ ^b	pK _a ^c	Caco-2 cell ^d (× 10 ⁶ cm/s)
6	2-F	1.215	5.419	10.484	5.928
8^c	2-Cl	1.432	4.790	10.485	7.503
9	3-Cl	1.456	4.974	10.487	3.869
10	4-Cl	1.420	5.378	10.487	3.869
13	4-Br	1.328	5.923	10.487	4.557
17	4-NO ₂	0.959	4.533	10.481	0.210
19	2,4-F ₂	1.143	4.823	10.483	3.313
20	3,4-Cl ₂	1.638	5.815	10.482	5.923
21	2-F,4-Cl	1.387	5.706	10.482	4.530
22	2-F,4-Br	1.347	6.146	10.482	5.254
Ponalrestat		1.678	4.441	3.360	0.464
Tolrestat		1.824	4.166	3.486	1.343

^a pIC₅₀ = −log(IC₅₀), where the IC₅₀ value is in μM.^b pED₅₀ = −log(ED₅₀), where the ED₅₀ value is in mol/kg/day.^c pK_a values were evaluated with the ACD/MedChem Advisor program.²⁹^d These values were calculated by using the PreADME program.³⁰^e The outlier.

above the physiological pH and maintain the neutral state suitable for the membrane transport. The high correlation between ED₅₀ and the Caco-2 cell permeability indicates that the hydrophobic character of spirosuccinimide type AR inhibitors warrants in vivo efficacy.

4. Materials and methods

4.1. Molecular structure

All spirosuccinimide congeners have a chiral center as shown in Table 1. We separately built the (*R*)-enantiomer and (*S*)-enantiomer molecular structures of 22 training compounds using Cerius2 program on Silicon Graphics Octane R10000 workstation.³¹ The lowest energy conformation of (*R*)-enantiomer was obtained by thorough conformational search and energy minimization. A set of conformers of a compound was generated by the grid search method and the molecular energy was minimized using the Merck Molecular Force Field.

The energy minimization was performed with the adopted basis Newton Raphson method until the root mean square of energy gradients reached 0.001 kcal/mol-Å. As the optimized (*R*)-enantiomer molecular structure of compound **22** was close to that of AS-3201 bound at the active site of AR,³² it was set to be the reference molecular frame. All other (*R*)-enantiomer structures were aligned to the reference structure through the flexible docking procedure.³² With the (*R*)-enantiomer structure determined, the initial structure of (*S*)-enantiomer was built by rotating the spirosuccinimide ring of the optimized and aligned (*R*)-enantiomer structure by 180° and was optimized by the same minimization protocol.

4.2. Molecular descriptors

The physicochemical descriptors for each enantiomer were calculated using the QSAR⁺ module of the Cerius2 program package.³¹ These descriptors include conformational, electronic, spatial, structural, thermodynamic, and molecular shape analysis descriptors. Atomic partial

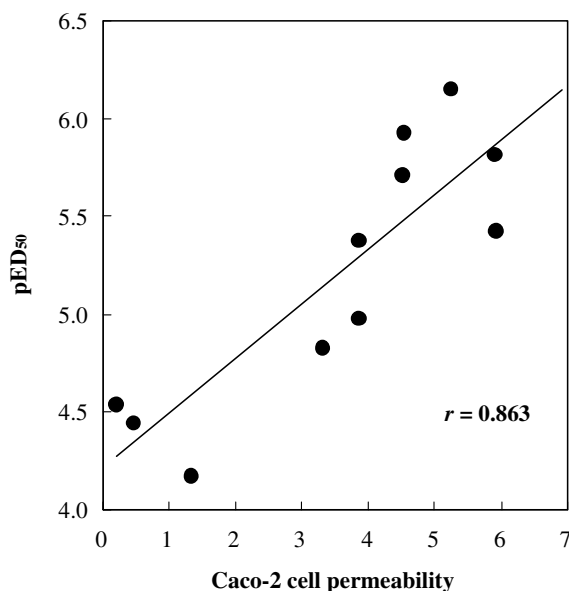


Figure 3. Correlation of observed pED₅₀ and the predicted Caco-2 cell permeability of 10 active spirosuccinimide AR inhibitors, ponalrestat and tolrestat. Compound **8** is excluded as the outlier.

charges were assigned to the optimized molecular structures using the charge equilibration method,³³ which were used to calculate the charge-dependent molecular descriptors such as dipole moment and polar surface area.

The Cerius2 program yielded two sets of descriptors: the *R*-descriptors for (*R*)-enantiomers and the *S*-descriptors for (*S*)-enantiomers. All the spirosuccinimide congeners were prepared as the racemic mixture, and subsequent biological activity assays were carried out with the racemates. We took the respective arithmetic mean value of *R*-descriptors and *S*-descriptors to generate the 'racemic' descriptor set, *RS*-descriptors. As shown previously for the in vitro QSAR model, the *RS*-descriptors were also checked to be suitable for in vivo QSAR analyses.

The *RS*-descriptor set included 78 physicochemical descriptors. All descriptor values and the activity data are normalized so that the coefficients of the resultant QSAR equation indicate the contribution of the descriptor to the activity. In order to remove interdependency among descriptors, we closely examined the correlation matrix of total 78 descriptors and searched for pairs of descriptors with the intercorrelation coefficient above 0.7. From the pair of highly correlated descriptors, we removed the one less correlated with the dependent variable, i.e., the in vivo percent inhibition activity. The percent inhibition and pED₅₀ strongly correlate with the correlation coefficient of 0.835 for the 10 active compounds used for the ED₅₀ assay. Thus, the screened set of *RS*-descriptors was used for both the percent inhibition and pED₅₀ QSAR analyses.

4.3. QSAR analyses

We employed the GFA technique²⁵ to derive the set of QSAR equations for the in vivo activities. The GFA

algorithm allows the construction of high-quality predictive models by evolving random initial models using a genetic algorithm.²⁵ Both linear and quadratic terms of descriptors were allowed in the independent variable set. The constructed models were evaluated with the *LOF* score and the model with the lowest *LOF* score was selected as the best one of the generation. We performed the GFA procedure with 20,000 genetic operations. A new term was added with the probability of 50%. The resultant QSAR equations were evaluated with the regression coefficient and the *LOF* score. The sequential *F* test was also performed to check if an additional term improved the QSAR model.

References and notes

- Laakso, M. *Diabetes* **1999**, *48*, 937.
- El-Kabbani, O.; Ruiz, F.; Darmanin, C.; Chung, R. P. T. *Cell. Mol. Life Sci.* **2004**, *61*, 750.
- Kador, P. F. *Med. Res. Rev.* **1988**, *8*, 325.
- Tomlinson, D. R.; Stevens, E. J.; Diemel, L. T. *Trends Pharmacol. Sci.* **1994**, *15*, 293.
- Lee, A. Y. W.; Chung, S. S. M. *FASEB* **1999**, *13*, 23.
- Nishimura, C. Y. *Pharmacol. Rev.* **1998**, *50*, 21.
- Yamagishi, M.; Yamada, Y.; Ozaki, K.; Asao, M.; Shimizu, R.; Suzuki, M.; Matsumoto, M.; Matsuoka, Y.; Matsumoto, K. *J. Med. Chem.* **1992**, *35*, 2085.
- Costantino, L.; Rastelli, G.; Vescovini, K.; Cignarella, G.; Vianello, P.; Del Corso, A.; Cappiello, M.; Mura, U.; Barlocco, D. *J. Med. Chem.* **1996**, *39*, 4396.
- Kotani, T.; Nagaki, Y.; Ishii, A.; Konishi, Y.; Yago, H.; Suehiro, S.; Okukado, N.; Okamoto, K. *J. Med. Chem.* **1997**, *40*, 684.
- Costantino, L.; Rastelli, G.; Gamberini, M. C.; Vinson, J. A.; Bose, P.; Iannone, A.; Staffieri, M.; Antolini, L.; Del Corso, A.; Mura, U.; Albasini, A. *J. Med. Chem.* **1999**, *42*, 1881.
- Costantino, L.; Rastelli, G.; Gamberini, M. C.; Giovannoni, M. P.; Piaz, V. D.; Vianello, P.; Barlocco, D. *J. Med. Chem.* **1999**, *42*, 1894.
- Fresneau, P.; Cussac, M.; Morand, J. M.; Szymanski, B.; Tranqui, D.; Leclerc, G. *J. Med. Chem.* **1998**, *41*, 4706.
- Nicolaou, I.; Zika, C.; Demopoulos, V. J. *J. Med. Chem.* **2004**, *47*, 2706.
- Settimo, F. D.; Primofiore, G.; Settimo, A. D.; Motta, C. L.; Simorini, F.; Novellino, E.; Greco, G.; Lavecchia, A.; Boldrin, E. *J. Med. Chem.* **2003**, *46*, 1419.
- Pau, A.; Asproni, B.; Boatto, G.; Grella, G. E.; Caprariis, P. D.; Costantino, L.; Pinna, G. A. *Eur. J. Pharm. Sci.* **2004**, *21*, 545.
- Bruno, G.; Costantino, L.; Curinga, C.; Maccari, R.; Monforte, F.; Nicolò, F.; Ottanà, R.; Vigorita, M. G. *Bioorg. Med. Chem.* **2002**, *10*, 1077.
- Costantino, L.; Corso, A. D.; Rastelli, G.; Petrash, J. M.; Mura, U. *Eur. J. Med. Chem.* **2001**, *36*, 697.
- Pfeifer, M. A.; Schumer, M. P.; Gelber, D. A. *Diabetes* **1997**, *46*, S82.
- Costantino, L.; Rastelli, G.; Vianello, P.; Cignarella, G.; Barlocco, D. *Med. Res. Rev.* **1999**, *19*, 3.
- Yasuda, H.; Terada, M.; Maeda, K.; Kogawa, S.; Sanada, M.; Haneda, M.; Kashiwagi, A.; Kikkawa, R. *Prog. Neurobiol.* **2003**, *69*, 229.
- Asano, T.; Saito, Y.; Kawakami, M.; Yamada, N. *J. Diabetes Complicat.* **2002**, *16*, 133.
- Negoro, T.; Murata, M.; Ueda, S.; Fujitani, B.; Ono, Y.; Kuromiya, A.; Komiya, M.; Suzuki, K.; Matsumoto, J. I. *J. Med. Chem.* **1998**, *41*, 4118.

23. Kim, J. R.; Won, Y. *Bull. Korean Chem. Soc.* **2004**, 25, 1874.
24. Ko, K.; Won, Y. *Bioorg. Med. Chem.* **2005**, 13, 1445.
25. Rogers, D.; Hopfinger, A. J. *J. Chem. Inf. Comput. Sci.* **1994**, 34, 854.
26. Stanton, D. T.; Jurs, P. C. *Anal. Chem.* **1990**, 62, 2323.
27. Ghose, A. K.; Crippen, G. M. *J. Comput. Chem.* **1986**, 7, 565.
28. Rohrbaugh, R. H.; Jurs, P. C. *Anal. Chim. Acta* **1987**, 199, 99.
29. ACD/MedChem Advisor, version 8.0; Advanced Chemistry Development Inc.: Toronto, Canada.
30. PreADME; Research Institute of Bioinformatics & Molecular Design: Seoul, Korea.
31. Cerius2, version 4.6; Accelrys Inc.: 6985 Scranton Road, San Diego, CA, USA.
32. Kurono, M.; Fujiwara, I.; Yoshida, K. *Biochemistry* **2001**, 40, 8216.
33. Rappé, A. K.; Goddard, W. A., III *J. Phys. Chem.* **1991**, 95, 3358.