—Article—

Modeling of Autocatalytic Hydrolysis of Adefovir Dipivoxil in Solid Formulations

Ying DONG,[a,b] Yan ZHANG,[a,b] Bingren XIANG,*[,b]
Haishan DENG,[c] and Jingfang WU[b]

[a]*Department of Organic Chemistry,* [b]*Center for Instrumental Analysis, China Pharmaceutical University, 24 Tongjia Xiang, 210009 Nanjing, Jiangsu, China, and* [c]*Key Laboratory for TCM Formulae Research of Jiangsu, College of Pharmacy, Nanjing University of Chinese Medicine, 138 Xianlin Road, 210046 Nanjing, Jiangsu, China*

(Received November 2, 2010; Accepted December 28, 2010)

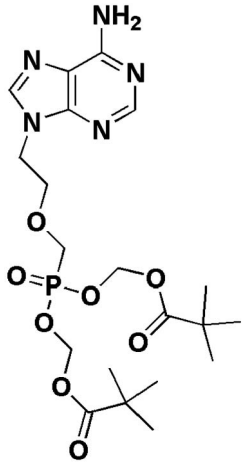
The stability and hydrolysis kinetics of a phosphate prodrug, adefovir dipivoxil, in solid formulations were studied. The stability relationship between five solid formulations was explored. An autocatalytic mechanism for hydrolysis could be proposed according to the kinetic behavior which fits the Prout-Tompkins model well. For the classical kinetic models could hardly describe and predict the hydrolysis kinetics of adefovir dipivoxil in solid formulations accurately when the temperature is high, a feedforward multilayer perceptron (MLP) neural network was constructed to model the hydrolysis kinetics. The build-in approaches in Weka, such as lazy classifiers and rule-based learners (*IBk*, *KStar*, *DecisionTable* and *M5Rules*), were used to verify the performance of MLP. The predictability of the models was evaluated by 10-fold cross-validation and an external test set. It reveals that MLP should be of general applicability proposing an alternative efficient way to model and predict autocatalytic hydrolysis kinetics for phosphate prodrugs.

Key words——adefovir dipivoxil; hydrolysis kinetics; autocatalysis; kinetic modeling; multilayer perceptron

INTRODUCTION

Phosphate esters play a significant role in the physiology of cells and therefore are fundamental to any organism. For the prodrugs of biologically active phosphates, bioactive protecting groups are commonly used as enormously powerful tool to increase bioavailability or to generally help deliver drugs of to cells. Ideally, these groups should be cleaved by intracellular enzymes, but stable outside cells. As known to all, a thorough understanding of chemical and physical stability of drugs and dosage forms is critical in the development and evaluation of pharmaceuticals. Although research on enzymatic hydrolysis of the phosphate prodrugs *in vivo* has already been covered thoroughly,[1] studies are still insufficient to characterize stability and hydrolysis mechanism for the phosphate prodrugs *in vitro*, especially in dosage form, which are essential to molecular design and formulation optimization for prodrugs of biologically active phosphates. In this study, adefovir dipivoxil (Fig. 1), a phosphate prodrug with masking group of pivaloyloxymethyl (POM) which was often employed to mask the negative charges of phosphate,[1] was used as a molecular model for investigating the stability and hydrolysis rules of prodrugs with biologically active acyloxyalkyl in solid phase.

Isothermal kinetic data for hydrolysis reaction of adefovir dipivoxil in different commercially available solid formulations was evaluated. Five classical kinetic models, including order-based and mechanism-based models, were used to describe the progress of the reaction. The Prout-Tompkins model fits the ther-



Adefovir dipivoxil (bis-POM PMEA)

Fig. 1. Chemical Structure of Adefovir Dipivoxil (bis-POM PMEA)

*e-mail: xiangbr@cpu.edu.cn

mal degradation of adefovir dipivoxil in solid formulations well. Therefore, the autocatalytic mechanism for hydrolysis could be proposed according to the kinetic behavior, suggesting an acid-catalyzed pathway.

Lee T. T. and co-workers[2] previously described the distribution of hydrolysis products of adefovir dipivoxil in the solid state (Fig. 2). The available literatures on nonenzymatic phosphate hydrolysis mechanism of phosphate diether including computational studies are mainly focused on base-catalyzed hydrolysis in solution.[3–6] Very few studies have addressed the acid-catalyzed hydrolysis mechanism relative to the reaction shown in Fig. 2. Basic research on this issue may eventually lead a better understanding of the acid-catalyzed hydrolysis mechanism for phosphate prodrugs in dosage forms.

Generally, formulation stability studies are time consuming and expensive. In this sense, the application of accurate and predictive machine learning methods should be both necessary and profitable. Machine learning techniques, such as neural networks[7–10], genetic algorithms,[11–13] decision trees[14] and support vector machines,[15] could automatically learn to recognize complex kinetic patterns and make intelligent decisions based on raw data of chemical processes, and it could allow to establish relationships between reaction conditions and results which attracts the attention primarily. This paper describes applications of some machine learning tools for modelling the kinetics data of the above hydrolysis reaction in solid formulations. Predictability of the machine learning models was evaluated by cross-validation and an external test set. This part of study treats of feasibility of machine learning method in modeling and predicting hydrolysis kinetics for phosphate prodrugs.

EXPERIMENTAL

Materials and Methods

Materials Adefovir dipivoxil (Fig. 1), also named bis-POM PMEA, with CAS number of 142340–99–6 is the second oral agent approved by the U.S. Food and Drug Administration (FDA) in 2002 for the treatment of chronic hepatitis B in patients 12 years of age and older.[16] Five kinds of solid formulations of adefovir dipivoxil (Sample I–V) were employed in this experiment. The content per dose for all the samples is 10 mg. Sample I, II and III are plain tablets, samples IV and V are hard capsules. Sample V with the trade name of Mingzheng® was obtained from Chia-tai Tianqing Pharmaceutical Co., Ltd. (Jiangsu, China). The other medicines including imported product were purchased from the local pharmacy. The compounds of c1 and c2 (Fig. 2) were synthesized by Chia-tai Tianqing Pharmaceutical Co., Ltd. (Jiangsu, China), and their chemical purity were 99.6% and 99.5% respectively. Acetonitrile (HPLC grade) was purchased from Merck (Darmstadt, Germany). Acetic acid, ammonium acetate, potassium chloride and magnesium chloride were AR grade and purchased from Shanghai Reagent Company (Shanghai, China). Deionized water was used through out the study.

Kinetic Methods

Storage The tablets and capsules were taken out from the original packages, placed in clear vials (2 ml, 4 tablets or 2 capsules per vial) uncapped and stored in a glass desiccator. The relative humidity

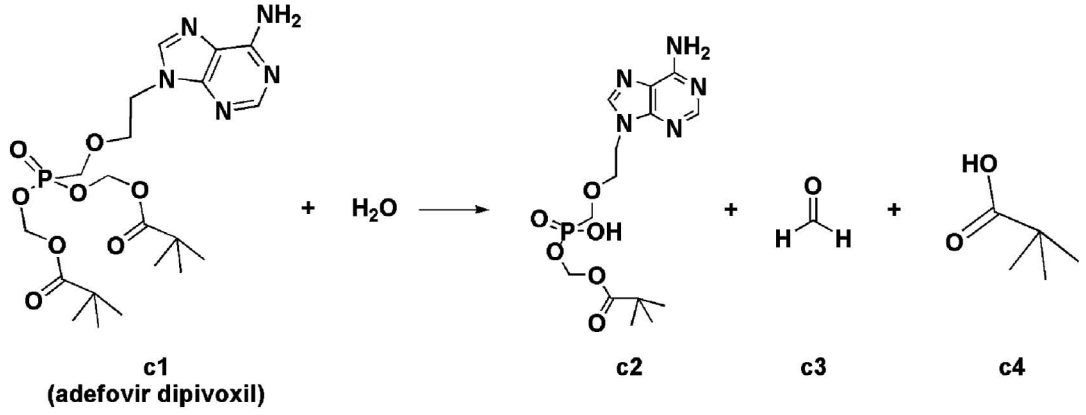


Fig. 2. Hydrolysis of Adefovir Dipivoxil in Solid State.

(RH) was controlled by placing saturated salt solutions in this closed desiccator. Put the desiccator in a temperature control cabinet with an accuracy of ±1 ℃. Chemical stability of each sample was examined under different conditions[17]: (a) 60℃/80% RH (KCl), (b) 60℃/29% RH ($MgCl_2$), (c) 75℃/79% RH (KCl) and (d) 75℃/27% RH ($MgCl_2$). One tablet or capsule was removed by random at each time point and stored in a refrigerator at −4℃ until they were analyzed. Repeat experiments ($n=3$) were carried out under identical conditions.

Pretreatment of the Samples In the pretreatment process, solution A consisting of acetonitrile and 0.025 mol/l ammonium acetate solution (adjusted pH to 6.0 with acetic acid) (35 : 65, v/v) was used. It was also employed in the HPLC and LC/MS assay as the mobile phase. One tablet or inclusion of a capsule was transferred to a 10-ml volumetric flask and 8 ml solution A was added. The mixture was vortex-mixed for 2 min and sonicated for 15 min. After equilibrating the mixture to room temperature, solution A was filled to the exact volume, and sonicated for another 15 min. The mixture was filtered through a disposable syringe filter unit (PTFE, 0.45 μm pore size) and the filtrate was collected for chromatographic analysis.

HPLC Assay Hydrolysis of adefovir dipivoxil in solid formulation was monitored by HPLC. The chromatographic system was the Shimadzu HPLC LC-10ATVP series (Tokyo, Japan) equipped with an LC-10 ATVP pump, a 7725 manual injector, an SPD-10 AVP detector, and an N2000 workstation (Zhejiang University, China). Elution was performed at room temperature in a reversed-phase Hanbon Kromasil C18 column (250 mm×4.6 mm I.D., 100 Å, 5 μm particle size, Hanbon Science & Technology Co., Ltd, China). The mobile phase, as mentioned above, consisted of acetonitrile and 0.025 mol /L ammonium acetate solution (adjusted pH to 6.0 with acetic acid) (35 : 65, v/v). The flow rate was 1.0 ml/min, the detection wavelength was 260 nm, and the injection volume was 10 μL. All solvents were filtered with filters (0.45 μm pore size) and degassed. The retention time of adefovir dipivoxil was found to be about 17 min. Baseline resolution was achieved for adefovir dipivoxil and its hydrolysis products. The correlation coefficient of the standard curve (concentration *vs*. peak area, between 2–1000 μg/ml) for adefovir dipivoxil was 0.9999.

Structural Characterization of Hydrolysis Product To verify the literature report,[2] chemical structure of hydrolysis product c2 in the dosage form was characterized by LC-MS. A Waters-2695 Alliance HPLC instrument (Waters Corporation, Milford, MA, USA) LC system with a 2996 photodiode array detector (DAD) was coupled with a micromass Q/TOF mass spectrometer with electrospray ionization. UV detection was achieved at 200–460 nm. A reversed-phase Hanbon Kromasil C18 column (250 mm× 4.6 mm I.D., 100 Å, 5 μm particle size, Hanbon Science & Technology Co., Ltd, China) was used. Solution A was used as mobile phase. The solvent flow rate was 0.9 ml/min and the column temperature was set at 30℃. The typical parameters are as follows: capillary voltage, 3 kV; sample cone voltage, 30 V; extraction cone voltage, 3 V; desolvation temp, 300 ℃; source temp, 100℃; full scan range, 100–1500 m/z. The hydrolysis product c2 was additionally characterized by comparing the retention time with that of an authentic sample.

Kinetic Models and Statistical Analysis Hydrolysis is one of the most common reactions seen with drug substances that contain ester functional group such as carboxylic ether and phosphate ether. In solid dosage forms, the amount of moisture is often relatively low. But hydrolysis of the ester can still be catalyzed and accelerated under specific circumstances. Carstensen and Attarchi used the Prout-Tompkins equation (Eq. (1))[18] to describe the sigmoid α-time curves obtained.

$$d\alpha/dt = k_B\alpha(1-\alpha) \quad (1)$$

Where α is the fractional extent of reaction, t is time, and k_B is the rate coefficient for this reaction.

It expresses the dependence of the rate on both the amount of reactant left and the amount of product formed, namely autocatalysis.[19] Theoretically, hydrolysis of ester could be catalyzed by the acid, which is one of the products formed by the reaction simultaneously. So autocatalysis in the hydrolysis reaction of ester is not only a kinetic feature, but also an useful criterion for exploring the reaction mechanism.[20,21]

In this study, adefovir dipivoxil content at time t (m_t) of each sample was measured. The built-in function median( ) of MATLAB 6.5 was used to return the median value of the data in the repeat experiments at time $t(m_t^*)$, which was employed in further processing.

The fractional extent of reaction (α) which defined

by Eq. (2) was calculated.

$$\alpha=\frac{m_0^*-m_t^*}{m_0^*} \quad (2)$$

Where m_0^* is the median value of initial weight of adefovir dipivoxil. Possible outliers in the data α was determined by Grubbs test ($p<0.05$) and eliminated. 155 data points (α values at time t) were left for modeling.

Hydrolysis of adefovir dipivoxil in solid formulation was characterized by the Prout-Tompkins equation (Eq. (1)) which on integration (see below) gives:

$$\ln\left(\frac{\alpha}{1-\alpha}\right)=kt+c \quad (3)$$

Where k is the rate coefficient for this reaction and c is the integration constant.

The other four equations were used as follows to model the kinetic data for comparison.

The modified Prout-Tompkins equation (Eq. (4))[22]:

$$\ln\left(\frac{\alpha}{1-\alpha}\right)=k\ln t+c' \quad (4)$$

Zero-order reaction model (Eq. (5)):

$$\alpha=kt+c'' \quad (5)$$

First-order reaction model (Eq. (6)):

$$\ln(1-\alpha)=-kt+c''' \quad (6)$$

Where k is the rate coefficient, c', c'' and c''' is the integration constant.

The Weibull equation (Eq. (7))[23]:

$$\ln\left(\ln\left(\frac{1}{1-\alpha}\right)\right)=\ln k+m\ln t \quad (7)$$

Where k is the rate coefficient and m is a constant.

The fractional extent of reaction (α) was plotted as a function of time (t). The rate coefficient (k) for Eq. (3)–Eq. (7) was obtained from the regression analysis using MATLAB 6.5 engineering software. Analysis of variance (ANOVA) was used to give a statistical test of the models.

Machine Learning Method

Construction of Multilayer Perceptron (MLP) MLP is a feedforward artificial neural network model that maps sets of input data onto a set of appropriate output. In this paper, the standard feedforward MLP neural networks were built. Weka 3.5.8[24] or called the Waikato Environment for Knowledge Analysis, a machine learning workbench which includes a framework in the form of Java class library, was employed. The settings were as follows: learning rate, 0.3; momentum, 0.2; the number of training epochs, 500. The *NominalToBinaryFilter* was turned on to help improve performance as there were nominal attributes in the data. The numerical variables were normalized with *normalizeAttributes*, and the nominal variables were normalized as well after they had been run through the *NominalToBinaryFilter*. The activation function is sigmoidal for the hidden neurons, and linear for the output neuron. Trials were made by adjusting the number of neurons in the hidden layers gradually while optimising the transfer function for the given input and output conditions.

155 data points (α values at time t) relevant to all the samples (I–V) under four different conditions constituted the original dataset. 30% of the data points were held out by random for testing and the remaining for training. The input variables were sample name (SN, nominal), Kelvin temperature (T, numerical), relative humidity (RH, numerical) and time (t, numerical). The output variable was α (numerical).

Model Validation In order to evaluate whether the learning algorithm of MLP was appropriate to this issue, the learning scheme on the *Classify* panel in Weka 3.6.0 was used as a model generator to build different models. Four kinds of classifier algorithms had been applied, they were two Lazy classifiers (*IBk* and *KStar*) and two Rules algorithms (*DecisionTable* and *M5Rules*). *IBk*[25] is a k-nearest-neighbor classifier that uses the Euclidean distance metric. *KStar*[26] is an instance-based classifier, the class of a test instance is based upon the class of those training instances similar to it, as determined by entropy-based distance function. *DecisionTable*[27] builds a decision table majority classifier and evaluates feature subsets using best-first search. *M5Rules*[28] generates a decision list for regression problems using separate-and-conquer. In each iteration, it builds a model tree using M5 and makes the ''best'' leaf into a rule. The status bar of classifier on the *Classify* panel was set as default.

The training set (108 data points) was treated by MLP and four additional classifier algorithms respectively. The overall predictive performance was detected using 10-fold cross-validation approach. An external test set (47 data points) was then used for external prediction to validate these models.

RESULTS AND DISCUSSION

Degradation Product Identities Chromatographic separation of the degraded adefovir dipivoxil

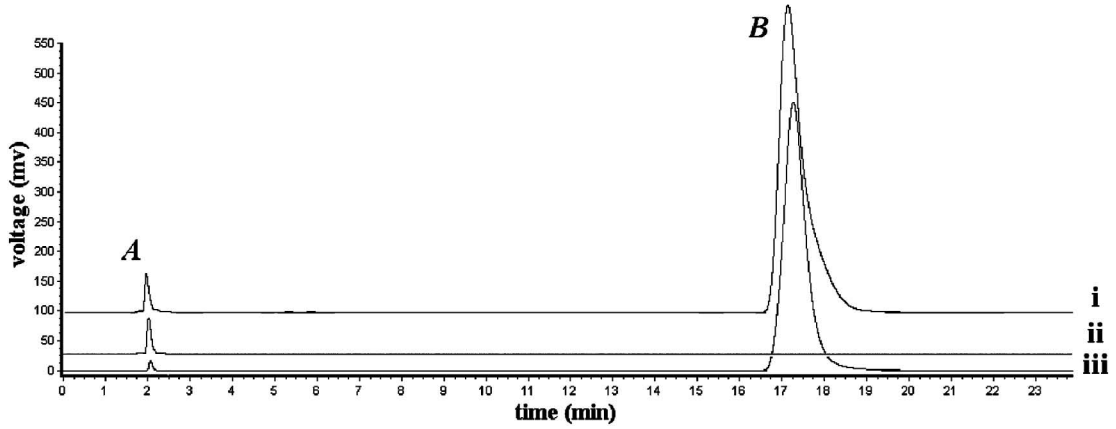


Fig. 3. HPLC Chromatograms of Adefovir Dipivoxil Sample and Its Hydrolysis
Product. i) Sample V degradation at 60℃/29% RH for 5 days. ii) Standard sample (c2). iii) Standard sample (c1).

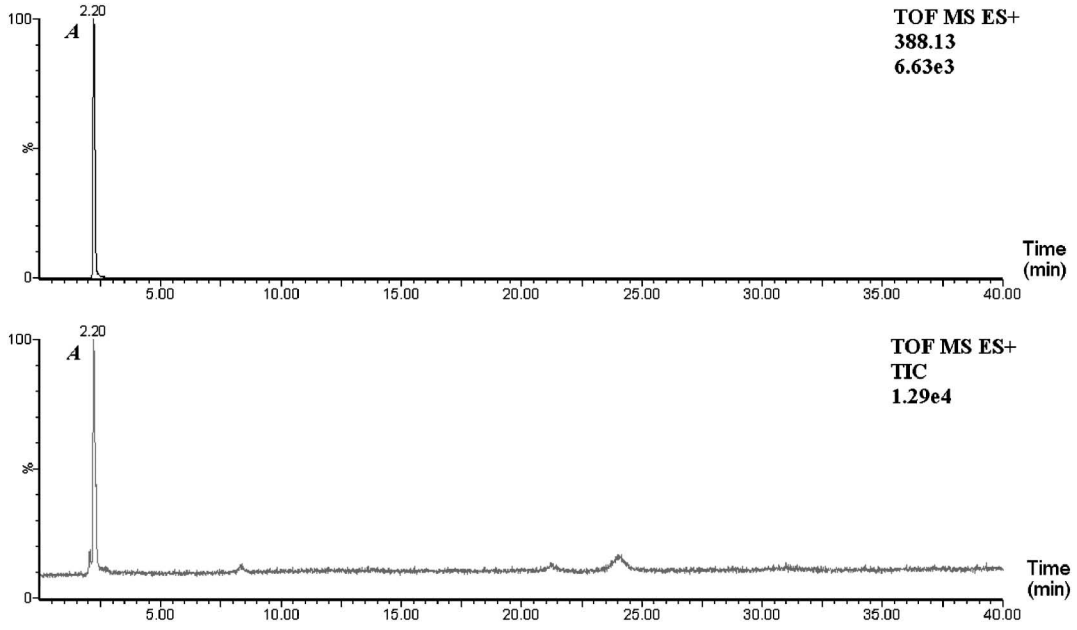


Fig. 4. Selected Ion Chromatography and Total Ion Chromatography.
The standard sample of c1 degradation at 60℃/80% RH for 7 days.

samples shows two major peaks (A and B, Fig. 3), which were identified by comparing retention time with that of standard samples (c2 and c1). The results confirm that peaks A and B correspond to c2 and c1, respectively.

In addition, peak A was found to have a mass of 387.13 (Fig. 4) which corresponding to compound c2. The proposed degradation route shown in Fig. 2 could be supported accordingly.

Stability Studies The conversion fraction (α) of adefovir dipivoxil in the samples was plotted as a function of time (Fig. 5). The shape of the curve is sigmoidal, as is common for solid-state decomposition. It is characterized by an induction period, a growth period and a decay period in turn. On the whole, the induction period is longer under relatively low temperature and (or) low RH, and meanwhile the slope of growth period becomes smaller.

With respect to the stability of different formulations, sample V should be best under these conditions. But the diversity of stabilities for different formulations is not relative to the difference between plain tables and hard capsules for the sample IV, as much a hard capsule as sample V, exhibits the worst stability. In other word, the excellent stability of sample V might owe to other features of the formulation,

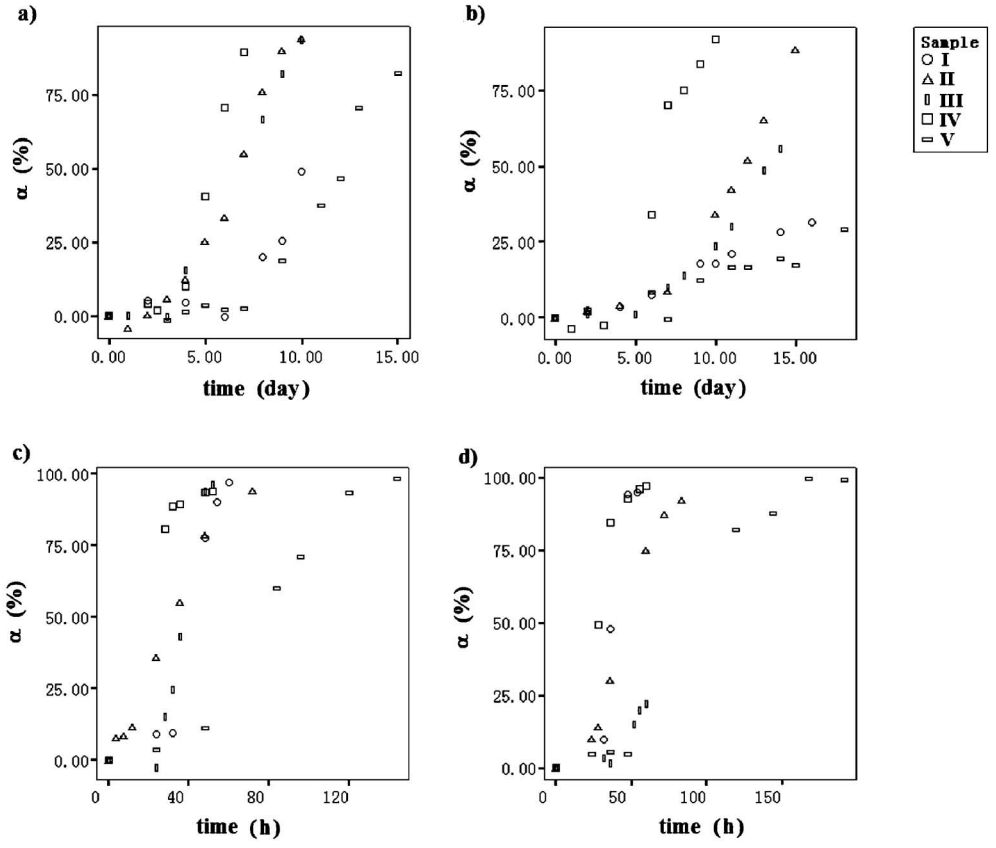


Fig. 5. α-Time Profile of Adefovir Dipivoxil in Solid Formulations Under Four Different Conditions.
a) 60℃/80% RH, b) 60℃/29% RH, c) 75℃/79% RH and d) 75℃/27% RH.

such as formulation design, hygroscopicity of the capsule, processing technique, crystal form of the active drug component, *etc*.

Figure. 5 also shows that there are some missing data points in the growth period at 75℃, which had been determined as outliers by Grubbs test ($p<0.05$) and eliminated. That is to say, the α value at time t in the repeats experiment changed remarkably. This phenomenon might account for the complex hydrolysis kinetics of adefovir dipivoxil in solid formulations under relatively high temperature.

Statistical Analysis on the Kinetic Data The results of linear regression analysis according to Eq. (3)–Eq. (7) have been listed in Table 1–Table 5. The Pearson product-moment correlation coefficient (typically denoted by R) was calculated, and a two-tailed correlation coefficient test was taken to check its significance. It should be noted that Eq. (3), Eq. (4) and Eq. (7) have lower limit when $\alpha=0$ or $t=0$, so the numbers of median value (N) are one point shorter than that of the others. In order to offer statistically good models, the process of linear regression analysis was neglected when $N<6$.

Figure. 6 summarizes the statistical results of correlation coefficient test and F test in linear regression. It is shown that 80% of the data groups could fit Eq. (3) well (one data group refers to a set of α and t for a specific sample under a given condition), *i.e.*, $\ln(\alpha/1-\alpha)$ is significantly related to t ($p<0.001$). The performance of Weibull equation (Eq. (7)) is a little poorer than that of Eq. (3).

As we mentioned above, the Prout-Tompkins equation (Eq. (3)) is a mechanism-based model which indicates an autocatalysis process. Theoretically, the kinetic behavior in these experiments could be explained through the acid autocatalysis that occurs in solid state hydrolysis of adefovir dipivoxil formulations. Figure. 2 shows there are two acid hydrolysis products (c2 and c4) that might catalyze the reaction. The real reaction mechanism could be extremely complicated, considering the effects of temperature, humidity, formulation, *etc*. Although this hydrolysis process was, in general, dominated by the Prout-Tompkins model, other models (such as Weibull

Table 1. The Results of Linear Regression Analysis According to Eq. (3)

Condition [A)]	Sample	Number of median value (N)	R	R square	F	Sig.	Correlation coefficient test [B)]	F test [B)]
a)	I	6	0.960	0.921	46.503	0.002	○	*
	II	10	0.998	0.995	1476.710	0.000	**	**
	III	6	0.984	0.968	120.274	0.000	**	**
	IV	6	0.955	0.911	41.163	0.003	○	*
	V	10	0.962	0.926	100.680	0.000	**	**
b)	I	8	0.968	0.936	88.412	0.000	**	**
	II	8	0.987	0.975	233.779	0.000	**	**
	III	8	0.996	0.993	821.629	0.000	**	**
	IV	8	0.991	0.981	210.589	0.000	**	**
	V	9	0.978	0.957	157.468	0.000	**	**
c)	I	5	—	—	—	—	—	—
	II	7	0.995	0.989	452.252	0.000	**	**
	III	6	0.959	0.919	45.515	0.003	○	*
	IV	5	—	—	—	—	—	—
	V	6	0.999	0.998	1837.525	0.000	**	**
d)	I	4	—	—	—	—	—	—
	II	6	0.994	0.987	308.175	0.000	**	**
	III	5	—	—	—	—	—	—
	IV	5	—	—	—	—	—	—
	V	7	0.969	0.939	76.789	0.000	**	**

Dependent Variable: $\ln(\alpha/1-\alpha)$, Independent Variable: t. [A)] a) 60℃/80% RH, b) 60℃/29% RH, c) 75℃/79% RH and d) 75℃/27% RH. [B)] ** denotes the variables are significantly related ($p<0.001$), * denotes related ($p<0.01$), and ○ denotes not related ($p>0.01$).

Table 2. The Results of Linear Regression Analysis According to Eq. (4)

Condition [A)]	Sample	Number of median value (N)	R	R square	F	Sig.	Correlation coefficient test [B)]	F test [B)]
a)	I	6	0.892	0.796	15.617	0.017	○	○
	II	10	0.969	0.939	108.214	0.000	**	**
	III	6	0.975	0.950	76.044	0.001	**	**
	IV	6	0.938	0.880	29.425	0.006	○	*
	V	10	0.969	0.939	123.088	0.000	**	**
b)	I	8	0.984	0.969	185.287	0.000	**	**
	II	8	0.921	0.847	33.336	0.001	*	*
	III	8	0.947	0.897	52.356	0.000	**	**
	IV	8	0.985	0.969	127.107	0.000	**	**
	V	9	0.951	0.904	66.146	0.000	**	**
c)	I	5	—	—	—	—	—	—
	II	7	0.947	0.897	43.726	0.001	*	*
	III	6	0.930	0.864	25.477	0.007	○	*
	IV	5	—	—	—	—	—	—
	V	6	0.963	0.928	51.581	0.002	○	*
d)	I	4	—	—	—	—	—	—
	II	6	0.999	0.999	2927.940	0.000	**	**
	III	5	—	—	—	—	—	—
	IV	5	—	—	—	—	—	—
	V	7	0.921	0.848	27.878	0.003	*	*

Dependent Variable: $\ln(\alpha/1-\alpha)$, Independent Variable: $\ln(t)$. [A)] a) 60℃/80% RH, b) 60℃/29% RH, c) 75℃/79% RH and d) 75℃/27% RH. [B)] ** denotes the variables are significantly related ($p<0.001$), * denotes related ($p<0.01$), and ○ denotes not related ($p>0.01$).

Table 3. The Results of Linear Regression Analysis According to Eq. (5)

Condition [A)]	Sample	Number of median value (N)	R	R square	F	Sig.	Correlation coefficient test [B)]	F test [B)]
a)	I	7	0.808	0.652	9.372	0.028	○	○
	II	11	0.962	0.925	110.547	0.000	**	**
	III	7	0.968	0.937	74.710	0.000	**	**
	IV	7	0.903	0.816	22.203	0.005	*	*
	V	11	0.912	0.831	44.224	0.000	**	**
b)	I	9	0.989	0.977	299.703	0.000	**	**
	II	9	0.937	0.878	50.525	0.000	**	**
	III	9	0.923	0.852	40.241	0.000	**	**
	IV	9	0.963	0.927	89.390	0.000	**	**
	V	10	0.926	0.857	47.996	0.000	**	**
c)	I	6	0.912	0.832	19.796	0.011	○	○
	II	8	0.987	0.974	227.650	0.000	**	**
	III	7	0.867	0.752	15.152	0.011	○	○
	IV	6	0.916	0.839	20.886	0.010	○	○
	V	7	0.975	0.950	94.998	0.000	**	**
d)	I	5	—	—	—	—	—	—
	II	7	0.973	0.947	89.568	0.000	**	**
	III	6	0.867	0.752	12.116	0.025	○	○
	IV	6	0.965	0.931	54.122	0.002	○	*
	V	8	0.972	0.946	104.428	0.000	**	**

Dependent Variable: α, Independent Variable: t. [A)] a) 60℃/80% RH, b) 60℃/29% RH, c) 75℃/79% RH and d) 75℃/27% RH. [B)] ** denotes the variables are significantly related ($p<0.001$), * denotes related ($p<0.01$), and ○ denotes not related ($p>0.01$).

Table 4. The Results of Linear Regression Analysis According to Eq. (6)

Condition [A)]	Sample	Number of median value (N)	R	R square	F	Sig.	Correlation coefficient test [B)]	F test [B)]
a)	I	7	0.778	0.605	7.658	0.039	○	○
	II	11	0.886	0.784	32.738	0.000	**	**
	III	7	0.915	0.838	25.860	0.004	*	*
	IV	7	0.837	0.701	11.729	0.019	○	○
	V	11	0.853	0.727	24.001	0.001	**	**
b)	I	9	0.986	0.972	242.312	0.000	**	**
	II	9	0.824	0.678	14.759	0.006	*	*
	III	9	0.919	0.844	32.386	0.001	**	**
	IV	9	0.930	0.866	45.076	0.000	**	**
	V	10	0.926	0.857	47.875	0.000	**	**
c)	I	6	0.864	0.746	11.752	0.027	○	○
	II	8	0.973	0.947	106.186	0.000	**	**
	III	7	0.800	0.641	8.916	0.031	○	○
	IV	6	0.986	0.972	137.540	0.000	**	**
	V	7	0.915	0.838	25.772	0.004	*	*
d)	I	5	—	—	—	—	—	—
	II	7	0.955	0.911	51.479	0.001	**	**
	III	6	0.861	0.742	11.480	0.028	○	○
	IV	6	0.966	0.933	55.479	0.002	○	*
	V	8	0.921	0.849	33.694	0.001	*	*

Dependent Variable: $\ln(1-\alpha)$, Independent Variable: t. [A)] a) 60℃/80% RH, b) 60℃/29% RH, c) 75℃/79% RH and d) 75℃/27% RH. [B)] ** denotes the variables are significantly related ($p<0.001$), * denotes related ($p<0.01$), and ○ denotes not related ($p>0.01$).

Table 5. The Results of Linear Regression Analysis According to Eq. (7)

Condition [A)]	Sample	Number of median value (N)	R	R square	F	Sig.	Correlation coefficient test [B)]	F test [B)]
a)	I	6	0.908	0.824	18.725	0.012	○	○
	II	10	0.993	0.987	528.132	0.000	**	**
	III	6	0.995	0.991	439.606	0.000	**	**
	IV	6	0.935	0.874	27.722	0.006	○	*
	V	10	0.970	0.940	125.267	0.000	**	**
b)	I	8	0.985	0.970	193.646	0.000	**	**
	II	8	0.950	0.902	55.498	0.000	**	**
	III	8	0.957	0.915	64.858	0.000	**	**
	IV	8	0.995	0.989	370.755	0.000	**	**
	V	9	0.959	0.919	79.500	0.000	**	**
c)	I	5	—	—	—	—	—	—
	II	7	0.970	0.941	80.249	0.000	**	**
	III	6	0.928	0.862	24.886	0.008	○	*
	IV	5	—	—	—	—	—	—
	V	6	0.991	0.982	223.236	0.000	**	**
d)	I	4	—	—	—	—	—	—
	II	6	0.997	0.993	605.933	0.000	**	**
	III	5	—	—	—	—	—	—
	IV	5	—	—	—	—	—	—
	V	7	0.940	0.883	37.638	0.002	*	*

Dependent Variable: ln (ln $(1/1-\alpha)$), Independent Variable: ln (t). [A)] a) 60℃/80% RH, b) 60℃/29% RH, c) 75℃/79% RH and d) 75℃/27% RH. [B)] ** denotes the variables are significantly related ($p<0.001$), * denotes related ($p<0.01$), and ○ denotes not related ($p>0.01$).

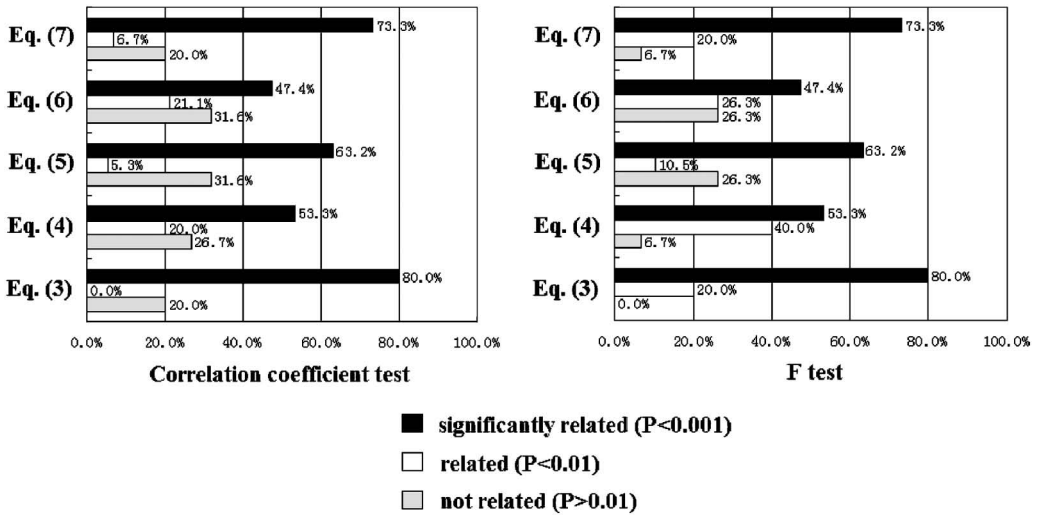


Fig. 6. The Comparison of Significance in Linear Regression for Evaluating Model Suitability.
The data groups with N<6 are neglected.

model) were also important to correctly describe the mechanism.

The observed reaction rate coefficient (k) was calculated for every data group according to linear regression result of Eq. (3) (Fig. 7). It shows that the stability relationship between formulations should be in accordance with the discussion introduced above basically, if the data groups of significantly related ($p<0.001$) are taken into account only. When the temperature is high (condition c and d), there are five

data groups whose N value is less than 6. That is to say, the α value at time t in the repeat experiments changed remarkably (these data points were eliminated in the data processing procedure). Therefore, the hydrolysis process becomes more complex when the temperature was high, and meanwhile, for the classical kinetic models, the rules of hydrolysis kinetics become inenarrable and unpredictable to some extent.

MLP Modeling The neural network of one hidden layer with four nodes was found to be superior in training the kinetic data. A scheme of the MLP model considered in this paper is given in Fig. 8. Several measures, which summarized in Table 6, were used to evaluate the success of numeric prediction during the training process. Although the *IBk* model has the highest correlation coefficient and the lowest errors while training, it could be estimated that this model would not perform accurately in prediction based on the results of 10-fold cross-validation. Furthermore, it is possible to say that the MLP, which has lower errors and higher correlation coefficient, should be the most predictable model.

The error rates on the independent test set are shown in Table 7. Figure 9 represents the plot of α predicted by MLP versus the experimental data during the training and test process. It can be observed that the MLP was good in predicting the hydrolysis kinetics of adefovir dipivoxil in solid formulations.

Comparing with the classical kinetic methods, the neural networks could model the full data as a whole and in consequence have no strictly limit on the N value for a specific data group. For that reason, neu-

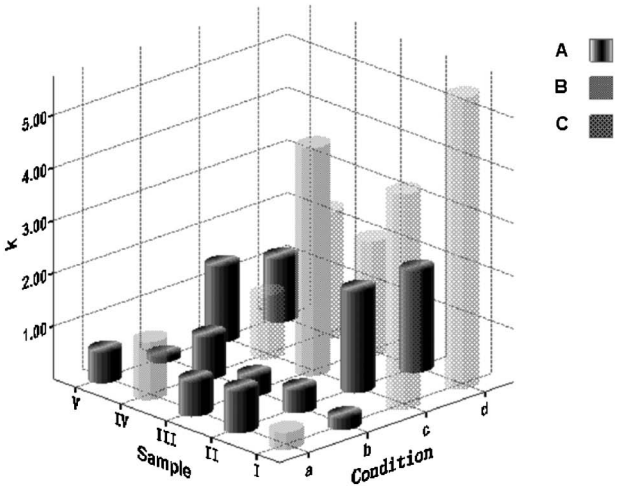


Fig. 7. Rate Coefficients (k, day^{-1}) for Every Data Group Obtained from the Prout-Tompkins Equation (Eq. (3))

a) 60℃/80% RH, b) 60℃/29% RH, c) 75℃/79% RH and d) 75℃/27% RH. A) The variables are significantly related ($p<0.001$) in correlation coefficient test and F test simultaneously. B) The variables are not related in correlation coefficient test ($p>0.01$) but related in F test ($p<0.01$). C) The number of median value (N) is less than 6.

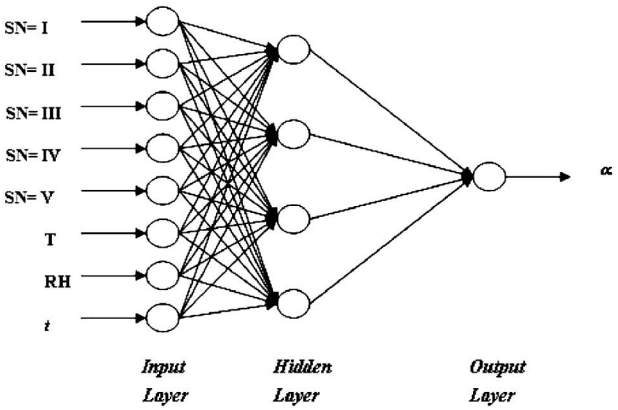


Fig. 8. Schematic Representation of the Three Layer Feed-forward Network Used in This Study

Table 6. Performance Measures on the Training Set

Test Mode	Model	Correlation coefficient	Mean absolute error	Root mean squared error	Relative absolute error (%)	Root relative squared error (%)
Full training	*MLP*	0.9696	6.1700	9.1851	17.8377	24.4688
	IBk	1.0000	0.0000	0.0000	0.0000	0.0000
	KStar	0.9810	7.3817	9.5538	21.3406	25.4510
	DecisionTable	0.7825	16.8392	23.3738	48.6826	62.2667
	M5Rules	0.9388	9.3762	13.1401	27.1070	35.0047
10-fold cross-validation	*MLP*	0.8945	12.0180	17.2998	34.5965	45.8542
	IBk	0.8481	12.3925	20.2985	35.6746	53.8026
	KStar	0.8832	15.3384	19.2728	44.1552	51.0838
	DecisionTable	0.7189	19.1243	26.2123	55.0537	69.4776
	M5Rules	0.8310	15.5097	20.9380	44.6482	55.4977

Table 7. Performance Measures on the Test Set

Model	Correlation coefficient	Mean absolute error	Root mean squared error
MLP	0.9219	6.5987	12.0106
IBk	0.8710	9.4109	16.0458
KStar	0.8074	17.0930	21.0280
DecisionTable	0.5480	24.6108	32.9236
M5Rules	0.8513	13.2747	16.0885

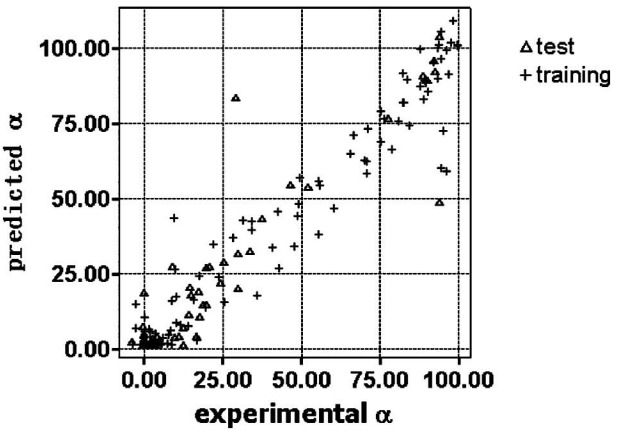


Fig. 9. Experimental Data and Predicted Kinetics Using MLP

ral networks have more tolerance for missing data and might be more useful for modeling a complex real-life process. But as a black box approach, this method could not extract and explain the intrinsic property and mechanism for the degradation of adefovir dipivoxil in solid formulations. Neural network-based kinetic modeling lay more emphasis on the practical world as opposed to the academic world.

CONCLUSION

In this study, stability and kinetic studies were made on hydrolysis of adefovir dipivoxil in solid formulations under controlled temperature and humidity. It reveals that relatively lower temperature and (or) lower RH are favorable for prolonging the induction period for solid-state decomposition. Sample V has good stability under experimental conditions. The stability of adefovir dipivoxil in solid formulations might owe to formulation design, hygroscopicity of the capsule, processing technique, crystal form of the active drug component, etc. The hydrolysis was found to proceed by autocatalytic mechanism according to the Prout-Tompkins model probably due to the acid products in existence. These findings will contribute to designing more stable prodrug molecules and developing more stable solid formulations for prodrugs that contain acyloxyalkyl masking groups.

Being an alternative for mechanism-based Prout-Tompkins models, multilayer perceptron was used to model properly the hydrolysis kinetics of adefovir dipivoxil. It provided a general method that suitable for analyzing complex kinetics in solid formulations with a relatively limited number of experiments and consequently reducing the cost of the stability experiment. And it is foreseeable that neural network models for degradation reactions could be rapidly obtained with less data if the well-trained models for similar reactions are available. Virtual screen for stable phosphate prodrugs could be realized supposing that plenty of the kinetics for the available prodrugs which undergo similar hydrolysis pathway have been collected to obtain high-quality predictions.

REFERENCES

1) Schultz C., *Bioorg. Med. Chem.*, **11**, 885–898 (2003).
2) Lee T. T., Munger J. D. Jr., Wu S., Krishnamurthy V. K., Abstracts of papers, AAPS Western Regional Meeting, San Francisco, April 1997, pp. 24–25.
3) Rosta E., Kamerlin S. C. L., *Warshel A., Biochemistry*, **47**, 3725–3735 (2008).
4) Chen X., Zhan C. G., *J. Phys. Chem. A*, **108**, 6407–6413 (2004).
5) Williams N. H., Wyman P., *Chem. Commun.*, **14**, 1268–1269 (2001).
6) Gorenstein D. G., Luxon B. A., Findlay J. B., *J. Am. Chem. Soc.*, **101**, 5869–5875 (1979).
7) Galván I. M., Zaldívar J. M., Hernández H., Molga E., *Comput. Chem. Eng.*, **20**, 1451–1465 (1996).
8) Baş D., Dudak F. C., Boyac1 İ. H., *J. Food Eng.*, **79**, 622–628 (2007).
9) Conesa J. A., Caballero J. A., Reyes-Labarta J. A., *J. Anal. Appl. Pyrolysis*, **71**, 343–352 (2004).
10) Serra J. M., Corma A., Argente E., Valero S., Botti V., *Appl. Catal. A: Gen.*, **254**, 133–145 (2003).
11) Edwards K., Edgar T. F., Manousiouthakis V. I., *Comput. Chem. Eng.*, **22**, 239–246 (1998).
12) Harris S. D., Elliott L., Ingham D. B., Pourkashanian M., Wilson C. W., *Comput. Meth.*

Appl. Mech. Eng., **190**, 1065–1090 (2000).
13) Majumdar S., Mitra K., *Chem. Eng. J.*, **100**, 109–118 (2004).
14) Varga T., Szeifert F., Abonyi J., *Eng. Appl. Artif. Intell.*, **22**, 569–578 (2009).
15) Li L. J., Su H. Y., Chu J., *Chin. J. Chem. Eng.*, **17**, 437–444 (2009).
16) Ayoub W. S., Keeffe E. B., *Aliment. Pharmacol. Ther.*, **28**, 167–177 (2008).
17) Nyqvist H., *Int. J. Pharm. Tech. Prod. Mfr.*, **4**, 47–48 (1983).
18) Prout E. G., Tompkins F. C., *Trans. Faraday Soc.*, **40**, 488–498 (1944).
19) Brown M. E., *Thermochim. Acta*, **300**, 93–106 (1997).
20) Lesutis H. P., Gläser R., Liotta C. L., Eckert C. A., *Chem. Comm.*, **20**, 2063–2064 (1999).
21) Krammer P., Vogel H., *J. Supercrit. Fluids*, **16**, 189–206 (2000).
22) Prout E. G., Tompkins F. C., *Trans. Faraday Soc.*, **42**, 468–472 (1946).
23) Yoshioka S., Stella V. J., ''Stability of Drugs and Dosage Forms,'' Kluwer Academic/Plenum Publishers, New York, 2000.
24) Witten I. H., Frank E., ''Data Mining: Practical Machine Learning Tools and Techniques,'' 2nd ed., Morgan Kaufmann, San Francisco, 2005.
25) Aha D. W., Kibler D., Albert M., *Mach. Learn.*, **6**, 37–66 (1991).
26) Cleary J. G., Trigg L. E., In Proceedings of the 12th International Conference on Machine Learning, Tahoe City, July 1995, pp. 108–114.
27) Kohavi R., In Proceedings of the 8th European Conference on Machine Learning, Crete, April 1995, pp. 174–189.
28) Holmes G., Hall M., Frank E., In Proceedings of the 12th Australian Joint Conference on Artificial Intelligence, Sydney, December 1999, pp. 1–12.