

ORIGINAL ARTICLE

A Measure Evaluating Relevance of a Validation Study of Alternatives to Animal Testing

Takashi Omori

Department of Biostatistics, Kyoto University, Japan

Abstract

Sensitivity, specificity and accuracy are well known measures for evaluating the relevance of an inter-laboratory validation study for alternative tests. It is not generally discussed that the measures are dependent on two determining factors: a set of chemicals and the number of laboratories. Furthermore, some alternative tests such as these for the phototoxicity test have an "Equivocal" category for judging the toxicity of chemicals. These facts have made it difficult to interpret the value of the measures.

Therefore, in this paper we propose new measures to evaluate the alternatives, which depend on a set of chemicals rather than on both factors, and can treat data which have "Equivocal" category. We also propose their confidence intervals, which are measures of their precision.

Key words: relevance, inter-laboratory validation study, sensitivity, specificity, accuracy, confidence interval

Introduction

Recently, due to an increasing social concern for animal welfare, a lot of alternative animal tests have been proposed, and in order to examine their feasibility and practicality various inter-laboratory validation studies have been conducted (e.g. Ray et al., 1994; Spielmann et al., 1998). Generally, the primary purpose of the validation study is to evaluate both the relevance and reliability of a proposed alternative test from the results of experiments using the alternative test (Balls et al., 1999). Sensitivity, specificity and accuracy are measures to determine the effectiveness of the alternative test when both the alternative and the animal tests have a binary classification for judging toxicity of chemicals, as "Positive" and "Negative". These are well known measures which have been widely used to evaluate the relevance of the alternative test in many validation studies (e.g. Balls et al., 1990; Roy et al., 1994; Spielmann et al., 1998).

However, two points should be taken into consideration concerning the interpretation of the summarized data from validation studies. The first point is that the values in the 2 by 2 table, which summarizes data, depend not only on a selected set of chemicals in the study but also on the number of participant laboratories. The other point is that a category for "Equivocal" produced from some alternative tests such as these for the phototoxicity test, which is neither a "Positive" nor "Negative" category, is often provided. For instance, the test guideline of the in vitro 3T3 NRU phototoxicity test states that '*a test substance with a PIF < 2 or an MPE < 0.1 predicts: "no phototoxicity". A PIF > 2 and < 5 or an MPE > 0.1 and < 0.15 predicts: "probable phototoxicity" and a PIF > 5 or an MPE > 0.15 predicts: "phototoxicity".*' where the PIF and the MPE are measurements of phototoxicity for the test (OECD, 2004). In this case, since there was a range suggesting similar performance when several cut-off points were examined, the

category "probable phototoxicity" as "Equivocal" was set (Peters and Holzhütter, 2002). Sugiyama, et al (1994) proposed a red blood cell hemolysis assay to predict phototoxicity of chemicals, and they classified photohemolysis into three categories, +, ± and -.

In this paper, we discuss the above two points for the measures, sensitivity, specificity and accuracy, and propose new measures for evaluating the relevance of an inter-laboratory validation study. We also construct an equation for their confidence intervals, which measure their precision of them (Altman, 2000a).

Methods

Definition for sensitivity, specificity and concordance

Table 1 shows a 2 by 2 table. Sensitivity is defined as the proportion of chemicals judged as positive by an alternative test in which the chemicals are identified as positive by an animal test. When data is summarized as in table 1, sensitivity is calculated by $a / (a + b)$. Specificity is defined as the proportion of chemicals judged as negative by the alternative test in which the chemicals are identified as negative by the animal test. The measure is $d / (c + d)$. Accuracy is defined as the proportion of a corresponding number of chemicals by the judgment of the alternative test in

which all the chemicals are identified by the animal test. The measure is obtained as $(a + d) / (a + b + c + d)$.

It is rarely noted that the values of these measures depend on the selected set of chemicals. If the toxicity of the selected chemicals in a validation study has only the strongest classes and the weakest classes, the values of these measures would be expected to be higher when the assessed alternative test has a good correlation to the targeted animal test. If the researchers conducting the validation study can select test chemicals before the experiments on the alternative test, they can control the measures. On the other hand, if they choose many middle class chemicals in the study, the measures may show an inferior result compared to our expectation. Even if the chemicals are selected by an external person not directly involved in the study, the values of these are dependent on the selected chemicals. Thus, we should interpret the values of these as conditional proportions dependent on the set of selected chemicals in the study.

Motivated data

Table 2 shows a typical form of data from a validation study. The symbols "P", "E" and "N" in the Table mean "Positive", "Equivocal" and "Negative" to be judged by *In vivo* test or the al-

Table 1. The 2 by 2 table.

		Animal test	
		Positive	Negative
Alternative test	Positive	a	c
	Negative	b	d
		a+b	c+d

Table 2. A motivated example of a inter-laboratory validation study.

Chemical	In vivo	Laboratory					
		a	b	c	d	e	f
A	P	P	E	E	P		
B	P	P	N	P	E		
C	N	P	P	P	P		
D	P	P	E			E	P
E	N	P	P			P	P
F	N	N	P			N	N
G	P			P	P	P	P
H	N			N	E	E	E
I	N			N	N	N	N

Symbols: P, positive; E, equivocal; N, negative

ternative test. This data is from an actual validation study conducted in Japan which has not been published yet. In the study, nine chemicals were tested by six laboratories. In order to meet an increasing demand for assessing test chemicals, the laboratories used the alternative test for as many chemicals as possible. However, due to time and financial constraints, all the laboratories did not experiment applying the alternative test for all the chemicals. In view of animal welfare, data from animal tests is usually obtained from some published articles and/or databases including data from past experiments; animal tests are rarely conducted in validation studies. Therefore there is usually only one result for each chemical. On the other hand, some results for each chemical in an alternative test are obtained from the inter-laboratory study.

When the measures, sensitivity, specificity and accuracy, are calculated, data, as in Table 2, is summarized by a 2 by 2 table, in which a result from a chemical in a laboratory for an alternative test corresponds to a result from using the same chemical in an animal test; total for four cells in the 2 by 2 table is 36 as is the case in Table.

Consideration of two points

Furthermore, in addition to the fact that the measures are a conditional proportion of a set of chemicals, we also have to consider that these depend on the number of laboratories conducting inter-laboratory validation studies. However, when data is summarized by a 2 by 2 table, as in Table 2, distinguishing between the two factors, the set of selected chemicals and the number of participant laboratories is overlooked. Then the interpretation of the value is difficult. For instance, the sensitivity from a laboratory which has examined ten positive chemicals is 100% when all the chemicals are judged positive. The sensitivity from the ten laboratories which examined a positive chemical is also 100% when all laboratories judge positive for the chemical. Should we regard both sensitivities as the same? Some people often use only the values of these measures from different validation studies without taking into consideration these factors, when they compare the alternatives.

The presence of an "Equivocal" category is another difficulty involved in interpreting the measures. Since these measures are based on the assumption that the results of both tests are expressed as binary categories, often data for "Equivocal" is artificially changed: these are eliminated from the numerator; data for "Equivocal"

cal" is relabeled as "Positive" (e.g. Sugiyama et al., 1994). The value of the measures depends on which treatment is used.

Proposed methods

We propose similar measures to sensitivity, specificity and accuracy, which take into consideration and deal with the previous two points.

Firstly, we consider the relationship between two factors; chemical and laboratory. Since several laboratories experiment using the alternative test for a same chemical in the inter-laboratory validation study, data from the validation study has a hierarchical structure between two factors. In the proposed methods, the factor of chemical becomes a basic unit.

Suppose y_{ij} is a variable to explain the result from an alternative test, and x_i is a variable to explain the result from an animal test, where subscript i and j mean the i th chemical ($i = 1, 2, \dots, n$) and the j th laboratory ($j = 1, 2, \dots, m_i$) respectively. The variable y_{ij} take 1 for the "Positive" result, 0 for the "Negative" and 0.5 for the "Equivocal", when the alternative test is experimented for the i th chemical in the j th laboratory. The variable x_i is 1 for the "Positive" result of the targeted animal test, and 0 for the "Negative" result. We initially define p_i as a proportion for the number of positive results in the i th chemical for the alternative test, that is

$$p_i = \sum_j y_{ij} / m_i. \quad (1)$$

As shown the appendix A, we can calculate the variance, $V(p_i)$, based on the assumption of trinomial distribution.

Using p_i , we also define q_i as

$$q_i = x_i p_i + (1 - x_i)(1 - p_i). \quad (2)$$

Note that q_i is a measure for the reliability of the i th chemical. The alternative test shows good reliability when the value of q_i is close to 1.

Finally, we define three measures which correspond to sensitivity, specificity and accuracy, using p_i , and call these measures Psn , Psp and Pac , respectively;

$$Psn = \sum_i x_i p_i / \sum_i x_i \quad (3)$$

$$Psp = \sum_i (1 - x_i)(1 - p_i) / \sum_i (1 - x_i),$$

and

$$Pac = \sum_i q_i / n = \left(\sum_i x_i p_i + \sum_i (1 - x_i)(1 - p_i) \right) / n \tag{5}$$

These measures are also consistent with sensitivity, specificity and accuracy which are based upon the number of tested chemicals, when each p_i becomes 1 or 0.

Using $V(p_i)$, we can also calculate the variance for the proposed measures, $V(Psn)$, $V(Psp)$ and $V(Pac)$ (Appendix B). Furthermore we can construct the 95% confidence interval as follows;

$$Psn \pm 1.96 \times \sqrt{V(Psn)},$$

$$Psp \pm 1.96 \times \sqrt{V(Psp)}, \text{ and}$$

$$Pac \pm 1.96 \times \sqrt{V(Pac)}.$$

Note that the proposed measures only depend on the set of chemicals in the validation study but not on the factor of laboratory.

Results

Table 3 displays a 2 by 2 table of the summarized data of Table 2, in which "Equivocal" for the alternative test is assigned to "Positive". Based on Table 3, sensitivity, specificity and accuracy become 93.7% (15/16), 40.0% (8/20) and 63.9% (23/36) respectively. When "Equivocal" is eliminated from the numerator, sensitivity and accuracy are 62.5%(10/16) and 50%(8/16) respectively.

Table 4 shows y_{ij} , x_i , m_i , p_i , q_i and $V(p_i)$ corresponding to Table 2. According to the table, the proposed measure, Psn , Psp and Pac are obtained as 78.1% (3.38/4), 47.5% (2.37/5) and 61.1%(5.75/9) respectively. The 95% confidence intervals of the proposed measures become

Table 3. The 2 by 2 table produced from Table 2. P: "Positive", E: "Equivocal".

		Animal test	
		Positive	Negative
Alternative Test	Positive	15 (P:10, E:5)	12 (P:9, E:3)
	Negative	1	8
		16	20

Table 4. Scores for data in Table 2, and p_i , q_i and $V(p_i)$.

Chemical	x_i (In vivo)	y_{ij} (Laboratory)						m_i	p_i	q_i	$V(p_i)$
		a	b	c	d	e	f				
A	1	1	0.5	0.5	1			4	0.75	0.75	0.25
B	1	1	0	1	0.5			4	0.625	0.625	0.688
C	0	1	1	1	1			4	1	0	0
D	1	1	0.5			0.5	1	4	0.75	0.75	0.25
E	0	1	1			1	1	4	1	0	0
F	0	0	1			0	0	4	0.25	0.75	0.75
G	1			1	1	1	1	4	1	1	0
H	0			0	0.5	0.5	0.5	4	0.375	0.625	0.188
I	0			0	0	0	0	4	0	1	0

24.7 to 100%, 9.5 to 85.5% and 31.2 to 91.1%, respectively.

When we set 0.5 for "Equivocal" and calculate the values of sensitivity, specificity and accuracy, these are consistent with P_{sn} , P_{sp} and P_{ac} , respectively. Though both measures are identical when the number of laboratories doing experiments for all the chemicals is the same, generally these return different values.

Discussion

There are several recommended statistical methods for data analysis when alternative tests are assessed (Festing, 2001). However, it is seldom that statistical methods for the inter-laboratory validation study have been developed. The proposed measures in this paper are for the inter-laboratory validation study.

Sensitivity, specificity and accuracy are commonly used in studies on diagnostic test studies in medicine (Altman, 1994). In these, the results of diagnostic tests between separate groups of patients with and without a target disease are summarized by a 2 by 2 table. One of the biggest differences compared to the situation for evaluating an alternative test is that the researcher is able to control the values of these measures by selecting the set of chemicals for the validation study. Our proposal measures also have the same feature.

This feature also affects the construction of the confidence intervals. Though the way to constructing of the confidence intervals for these measures is also well known (Altman, 2000b), it is assumed that each observation is followed by an independent and identical distribution. It would be difficult to make the assumption that experimental results from different chemicals have the same distribution because toxicity for chemicals selected in a study is usually widespread. However, the proposed measures solved the problem and we can construct their confidence intervals.

The confidence interval is an index to show the precision of measures, and usually it is affected by the number of data. Thus, $V(p_i)$ depends on the number of laboratories on the i th chemical, m_i ; $V(P_{sn})$, $V(P_{sp})$ and $V(P_{ac})$ also depend on the number of chemicals, n beside m_i . Therefore, if a validation study with a few laboratories is conducted, the confidence intervals of the proposed measures, P_{sn} , P_{sp} and P_{ac} , may become wider. Though this feature is a limitation of the proposed measure, we could obtain more reliable estimates

if there are replications for a chemical in a laboratory.

Acknowledgment

This study was supported by a grant from the Ministry of Health, Labour and Welfare.

References

- Altman, D. G. (2000a) Confidence intervals in practice, in *Statistics with confidence 2nd edition*, ed. by D. G. Altman, D. Machin, T. N. Bryant and M. J. Gardner, pp.6-14, BMJ.
- Altman, D. G. (2000b) Diagnostic tests, in *Statistics with confidence 2nd edition*, ed. by D. G. Altman, D. Machin, T. N. Bryant and M. J. Gardner, pp.105-119, BMJ.
- Altman, D. G. and Bland, J. M. (1994) Diagnostic tests 1: sensitivity and specificity, *BMJ*, **308**, 1552.
- Balls, M. and Fentem, J. H. (1999) The validation and acceptance of alternatives to animal testing, *Toxicology in Vitro*, **13**, 837-846.
- Balls, M., Blaauboer, Bas., Brusick, D., Frazier, J., Lamb, D., Pemberton, M., Reinhardt, C., Roberfroid, M., Rosenkranz, H., Schmid, B., Spielmann, H., Stammati, A.-L. and Walum, E. (1990) Report and recommendations of the CAAT/ERGATT workshop on the validation of toxicity test procedures, *ATLA*, **18**, 313-337.
- Peters, B. and Holzhütter, H.-G. (2002) In Vitro phototoxicity testing: Development and validation of a new concentration response analysis software and biostatistical analyses related to the use of various prediction models, *ATLA*, **30**, 415-432.
- Festing, M. F. (2001) Guidelines for the design and statistical analysis of experiments in papers submitted to ATLA, *ATLA*, **29**, 427-446.
- OECD (2004) OECD guideline for testing of chemicals:432 In Vitro 3T3 NRU phototoxicity test. Paris, France: Organisation for Economic Cooperation and Development.
- Roy, T. A., Saladdin, K. A. and Mackerer, C. R. (1994) Evaluation of the EYTEX system as a screen for eye irritancy of petroleum products, *Toxicology in Vitro*, **8**, 797-798.
- Spielmann, H., Balls, M., Dupuis, J., Pape, W. J., Pechovitch, G., de Silva, O., Holzhütter, H.-G., Clothier, R., Desolle, P., Gerberick, F., Liebsch, M., Lovell, W. W., Maurer, T., Pfannenbecker, U., Potthast, J. M., Csato, M., Sladowski, D., Steiling, W. and Brantom, P. (1998) The international EU/COLIPA In vitro phototoxicity validation study: Results of Phase II (Blind Trial). Part 1: The 3T3 NRU phototoxicity test, *Toxicology in Vitro*, **12**, 305-327.

Sugiyama, M., Itagaki, H., Hariya, T., Murakami, N. and Kato, S. (1994) In Vitro assay to predict phototoxicity of chemicals: (I) Red blood cell hemolysis assay, *AATEX*, 2, 183-191.
 Sugiyama, M., Mori, M., Hoya, M., Hirota, M. and Itagaki, H. (2002) A strategic approach for predicting phototoxicity of chemical ingredients, *AATEX*, 9, 29-39.

Corresponding author:
 Takashi Omori, Ph.D.
 Department of Biostatistics, Kyoto University,
 Yoshida-Konoecho, Sakyo-ku, Kyoto 606-8501,
 Japan
 Tel:+81-75-753-4482
 Fax:+81-75-753-4487
 E-mail: omori@pbh.med.kyoto-u.ac.jp

(Received: April 7, 2006 /
 Accepted: September 12, 2006)

Appendix A

Suppose y_{ij} is a score of the i th chemical in the j th laboratory, which takes 1 for "Positive", 0 for "Negative" and 0.5 for "Equivocal"; $r_i^{(N)}$, $r_i^{(E)}$ and $r_i^{(P)}$ are frequencies for "Negative", "Equivocal" and "Positive" of the i th chemical; m_i is the number of laboratories which experimented the i th chemical. We

assume that $(r_i^{(N)}, r_i^{(E)}, r_i^{(P)})$ follows a trinomial distribution with parameters, $m_i, \pi_i^{(N)}, \pi_i^{(E)}$ and $\pi_i^{(P)}$:
 $(r_i^{(N)}, r_i^{(E)}, r_i^{(P)}) \sim tri(m_i, \pi_i^{(N)}, \pi_i^{(E)}, \pi_i^{(P)})$, (A1)

where $\pi_i^{(N)} + \pi_i^{(E)} + \pi_i^{(P)} = 1$.

The sum of y_{ij} for the i th chemical can write use $(r_i^{(N)}, r_i^{(E)}, r_i^{(P)})$ is:

$$\sum_j y_{ij} = 0 \times r_i^{(N)} + 0.5 \times r_i^{(E)} + 1 \times r_i^{(P)}. \quad (A2)$$

Therefore, the expected value of $\sum_j y_{ij} / m_i$ is:

$$E\left[\sum_i y_i / m_i\right] = 0 \times \pi_i^{(N)} + .5 \times \pi_i^{(E)} + 1 \times \pi_i^{(P)} \\ = .5 \times \pi_i^{(E)} + 1 \times \pi_i^{(P)}. \quad (A3)$$

The variance is:

$$V\left[\sum_i y_i / m_i\right] \\ = 0 \times \frac{\pi_i^{(N)}(1 - \pi_i^{(N)})}{m_i} + .5 \times \frac{\pi_i^{(E)}(1 - \pi_i^{(E)})}{m_i} + 1 \times \frac{\pi_i^{(P)}(1 - \pi_i^{(P)})}{m_i} \\ - 2 \times \left(0 \times .5 \times \frac{\pi_i^{(N)}(1 - \pi_i^{(E)})}{m_i}\right) - 2 \times \left(.5 \times 1 \times \frac{\pi_i^{(E)}(1 - \pi_i^{(P)})}{m_i}\right) - 2 \times \left(0 \times 1 \times \frac{\pi_i^{(N)}(1 - \pi_i^{(P)})}{m_i}\right) \\ = (.5)^2 \times \frac{\pi_i^{(E)}(1 - \pi_i^{(E)})}{m_i} + \frac{\pi_i^{(P)}(1 - \pi_i^{(P)})}{m_i} - \frac{\pi_i^{(E)}(1 - \pi_i^{(P)})}{m_i}. \quad (A4)$$

Since we can't know the values of $\pi_i^{(N)}$, $\pi_i^{(E)}$ and $\pi_i^{(P)}$, we need to estimate these values as $\hat{\pi}_i^{(N)}$, $\hat{\pi}_i^{(E)}$ and $\hat{\pi}_i^{(P)}$. When the maximum likelihood method is used, the estimates are $\hat{\pi}_i^{(N)} = r_i^{(N)} / m_i$, $\hat{\pi}_i^{(E)} = r_i^{(E)} / m_i$ and $\hat{\pi}_i^{(P)} = r_i^{(P)} / m_i$. Then, by applying the estimates to (A3) and (A4), the values of p_i and $V(p_i)$ are:

$$p_i = .5 \times r_i^{(E)} / m_i + 1 \times r_i^{(P)} / m_i = \sum_j y_{ij} / m_i, \quad (A5)$$

$$V(p_i) = \frac{1}{4} \times \frac{r_i^{(E)}(m_i - r_i^{(E)})}{m_i} + \frac{r_i^{(P)}(m_i - r_i^{(P)})}{m_i} - \frac{r_i^{(E)}(m_i - r_i^{(P)})}{m_i}. \quad (A6)$$

Appendix B

Since p_i are independent for each other, using $V(p_i)$ the variances of P_{sn} , P_{sp} and P_{ac} , $V(P_{sn})$, $V(P_{sp})$ and $V(P_{ac})$, are obtained as:

$$\begin{aligned} V(P_{sn}) &= V\left(\frac{\sum_i x_i p_i}{\sum_i x_i}\right) \\ &= \frac{\sum_i x_i^2 V(p_i)}{\left(\sum_i x_i\right)^2} \\ &= \frac{\sum_i x_i V(p_i)}{\left(\sum_i x_i\right)^2}, \end{aligned} \tag{B1}$$

$$\begin{aligned} V(P_{sp}) &= V\left(\frac{\sum_i (1-x_i)(1-p_i)}{\sum_i (1-x_i)}\right) \\ &= \frac{\sum_i (1-x_i)^2 V(1-p_i)}{\left(\sum_i (1-x_i)\right)^2} \\ &= \frac{\sum_i (1-x_i) V(p_i)}{\left(\sum_i (1-x_i)\right)^2}, \text{ and} \end{aligned} \tag{B2}$$

$$\begin{aligned} V(P_{ac}) &= V\left(\frac{\sum_i x_i p_i + \sum_i (1-x_i)(1-p_i)}{n}\right) \\ &= \frac{\left\{\sum_i x_i V(p_i) + \sum_i (1-x_i) V(p_i)\right\}}{n^2} \\ &= \frac{\sum_i V(p_i)}{n^2}. \end{aligned} \tag{B3}$$

QUANTITATIVE MEASUREMENT OF SPLICED XBP1 mRNA AS AN INDICATOR OF ENDOPLASMIC RETICULUM STRESS

Morihiko HIROTA¹, Masato KITAGAKI¹, Hiroshi ITAGAKI¹ and Setsuya AIBA²

¹Shiseido Research Center, 2-11-1 Fukuura, Kanazawa-ku, Yokohama, Kanagawa 236-8643, Japan

²Department of Dermatology, Tohoku University of Medicine, 1-1 Seiryō-machi, Aoba-ku
Sendai, Miyagi 980-8574, Japan

(Received January 18, 2006; Accepted February 20, 2006)

ABSTRACT — The unfolded protein response (UPR) events triggered by the accumulation of unfolded protein in endoplasmic reticulum (ER) activate the three UPR signaling pathways mediated by IRE1, ATF6 and PERK. Spliced XBP1 mRNA induced by activated IRE1 is translated to the protein, a potent transcription factor that induces BiP expression. XBP1 is also induced by activated ATF6. It is thus thought to be an important marker reflecting both IRE1 and ATF6 signaling in response to ER stress. For quantitative measurement of XBP1 gene expression, it is important to distinguish between the spliced and non-spliced form of XBP1 mRNA. We have developed a new method to detect the spliced XBP1 mRNA by means of real-time PCR and we compared the result with measurements of the expression of the ER stress inducible gene BiP. A good correlation was found between spliced XBP1 expression and BiP expression. Thus, our method may be useful for simple and quantitative evaluation of ER stress.

KEY WORDS: Spliced XBP1, Quantitative measurement, THP-1, ER stress, BiP

INTRODUCTION

In eukaryotic cells, proteins translated from mRNA must be correctly folded and assembled in the endoplasmic reticulum (ER). If proteins are not folded correctly, they accumulate in the ER, and this accumulation of unfolded or misfolded proteins triggers the unfolded protein response (UPR) (Sidrauski *et al.*, 1998). In mammalian cells, there are three UPR signaling pathways, with three proteins (IRE1, ATF6 and PERK) as their UPR transducers and one master regulator (BiP) (Zhang and Kaufman, 2004). IRE1 (inositol-requiring enzyme 1) has two isoforms: IRE1 α and IRE1 β . IRE1 contains an N-terminal ER stress-sensing domain in the ER lumen, an ER transmembrane domain, and a serine/threonine kinase domain and a C-terminal endoribonuclease domain in the cytosol (Mori *et al.*, 1993; Shamu and Walter, 1996; Welihinda and Kaufman, 1996). Activating transcription factor 6 (ATF6) is a transcription factor with an N-terminal basic leucine zipper (b-ZIP) domain in the cytosol and a C-terminal ER luminal stress-sensing domain (Haze

et al., 1999). Protein kinase-like ER kinase (PERK) consists of an ER luminal stress-sensing domain and a cytosolic domain that phosphorylates the α subunit of eukaryotic translation initiation factor 2 (eIF2 α) (Shi *et al.*, 1998; Harding *et al.*, 1999; Liu *et al.*, 2000). BiP (GRP78: Glucose related protein 78) is a member of the Hsp70 chaperone family that plays critical roles in ER protein folding and quality control (Ellgaard and Helenius, 2003; Gething, 1999; Shen *et al.*, 2002). In unstressed cells, BiP binds to the luminal domain (LD) of IRE1, ATF6 and PERK. In response to accumulation of unfolded and misfolded proteins in the ER, BiP is released from these proteins. ATF6 is released from BiP association, transported to the Golgi apparatus and cleaved by S1P and S2P proteases to generate cytosolic domain of ATF6. The cytosolic fragment of ATF6 becomes functional as a transcription factor that binds the *cis*-acting ER stress response element (ERSE), the consensus site of which is CCAAT-N9-CCACG (Yoshida *et al.*, 1998). PERK and IRE1 released from BiP are dimerized and activated. Activated PERK inhibits mRNA translation by phosphorylating eIF2 α .

Activated IRE1 acquires RNase activity to initiate XBP1 (X-box DNA binding protein, UPR-specific b-ZIP transcription factor) mRNA splicing. Spliced XBP1 mRNA is translated to a potent transcription factor that binds to ERSE and the unfolding protein response element (UPRE), the consensus site of which is TGACGTGG/A (Yoshida *et al.*, 2001).

Recently, it has been suggested that ER stress may be involved in the pathogenesis the neurodegenerative diseases, such as Alzheimer's and Parkinson's diseases (Kudo *et al.*, 2002; Ryu *et al.*, 2002; Schroder and Kaufman, 2005) and diabetes (Harding and Ron, 2002). Although expression of BiP protein, ATF6 activation, splicing of XBP1 mRNA and phosphorylation of eIF-1 α are all potential biomarkers of ER stress, some of them are difficult to measure quantitatively. In this study, we have developed a new method to assay spliced XBP1 mRNA by using a real-time PCR system, and we show that the results obtained with this method correlate well with data for BiP.

MATERIALS AND METHODS

Chemicals

Dithiothreitol (DTT), tunicamycin (Tm) and thapsigargin (TG) were purchased from Wako Pure Chemicals (Osaka, Japan). Brefeldin A (BFA), A23187 were purchased from Sigma-Aldrich Corporation (St. Louis, MI). Dimethyl sulfoxide (DMSO) was purchased from Kanto Chemical (Tokyo, Japan).

Cell culture

THP-1 cells were obtained from the American Type Culture Collection (Rockville, MD). These cells were maintained in RPMI 1640 medium (Invitrogen life technologies, Carlsbad, CA) with 1% (v/v) antibiotic-antimycotic (Invitrogen life technologies), 10% FBS (v/v) (JRH Biosciences, Lenexa, KS) and 0.05 mM 2-mercaptoethanol (Invitrogen life technologies) at 37°C in a 5% CO₂ incubator. Cells were maintained to 0.1 – 0.5 × 10⁶ cells/mL by passage.

Total RNA isolation

THP-1 cells were seeded at 1 × 10⁶ cells/mL. Cells treated with various chemicals for 2 hr were collected by centrifugation (700 × g, 3 min, 4°C) and resuspended in 1 mL of Isogen reagent (Nippon Gene, Toyama, Japan), and total RNA was isolated as recommended by the manufacturer. Total RNA concentration and purity were determined by measuring OD₂₆₀ and OD_{260/280} ratio, respectively, on a NanoDrop spectro-

photometer (NanoDrop Technologies, Rockland, DE).

Primer sequences for semi-quantitative reverse transcriptase-polymerase chain reaction (RT-PCR) and real-time PCR

The expression of spliced XBP1 was measured by means of two PCR procedures (methods 1 and 2). Primer sequences for semi-quantitative RT-PCR and real-time PCR were as follows: XBP1 mRNA (method 1), <sense primer> 5'- CCTTG TAGTTGAGAAC-CAGG-3', <anti-sense primer> 5'-GGGCTTGG-TATATATGTGG-3'; XBP1 mRNA (method 2), <sense primer> 5'- GGTCTGCTGAGTCCGCAGCAGG -3', <anti-sense primer> 5'-GGGCTTGGTATATATGTGG-3'; BiP mRNA, <sense primer> 5'- CGAGGAGGAG-GACAAGAAGG -3', <anti-sense primer> 5'- CACCT-TGAACGGCAAGAACT-3'; GAPDH mRNA, <sense primer> 5'-GAAGGTGAAGGTCGGAGTC-3' <anti-sense primer> 5'-GAAGATGGTGATGGGATTTC-3'. For measurement of spliced XBP1 by method 2, we used a sense primer designed to span the 26 base pair (bp) intron, and thus to anneal only the spliced XBP1 mRNA.

Semi-quantitative RT-PCR

Total RNA samples were reverse-transcribed with a GeneAmp RNA PCR kit (Applied Biosystems, Foster City, CA). Briefly, 1 μ g of RNA was added to a tube containing 5 mM MgCl₂, 1 mM each of dNTP (A, T, G, C), 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 2.5 μ M random hexamer, 1 U of RNase inhibitor, and 2.5 U of murine leukemia virus (MuLV) reverse transcriptase. The mixture was incubated at 42°C for 60 min, heated to 95°C for 5 min, and placed on ice until used for PCR. The double-stranded cDNA was synthesized from single-stranded cDNA by PCR in a total volume of 20 μ L containing 7 pmol each of specific sense and anti-sense primers, 1.5 mM MgCl₂, 2.5 mM each of dNTP (A, T, G, C), 50 mM KCl, 10 mM Tris-HCl (pH 8.3), 0.5 units of AmpliTaq DNA polymerase (Applied Biosystems). For measurement of spliced XBP1 mRNA in method 1, XBP1 double-stranded cDNA was synthesized under the following thermal cycling conditions: 94°C 5 min, 95°C 30 sec - 55°C 30 sec - 72°C 30 sec 2 cycles. Then, 7.5 U of restriction enzyme *Pst*I (TaKaRa Bio, Shiga, Japan) was added to the reaction mixture for 1 hour to digest the double-stranded cDNA of unspliced XBP1. The remaining spliced XBP1 cDNA was amplified by PCR (95°C 5 min, 95°C 30 sec - 55°C 30 sec - 72°C 30 sec 28 cycles, 72°C 7 min). For measurement of spliced

Quantitative measurement of spliced XBP1 mRNA.

XBP1 mRNA in method 2, the PCR conditions were 95°C 5 min, 95°C 30 sec - 60°C 30 sec - 72°C 30 sec 30 cycles, 72°C 7 min. GAPDH mRNA expression was used as an internal control (measured under the same PCR conditions as used for XBP1 mRNA, methods 1 and 2). PCR products were analyzed by electrophoresis through 2.0% agarose gel, and their identity was checked by DNA sequencing.

Cloning of spliced and unspliced XBP1 cDNAs

Spliced and unspliced XBP1 cDNAs were amplified by semi-quantitative RT-PCR from total RNA in non-treated THP-1 and DTT-treated THP-1. Each amplified cDNA was analyzed by electrophoresis through 2.0% agarose gel, recovered from the gel and ligated into a pGEM[®]-T Easy vector (Promega Corporation, Madison, WI). The plasmid DNA from transformed *E. coli* JM109 cells was prepared using a QIAGEN[®] Plasmid Mini Kit (QIAGEN, K.K., Tokyo, Japan) and checked by DNA sequencing. The DNA concentration and purity were determined by measuring OD₂₆₀ and OD_{260/280} ratio, respectively, on a NanoDrop spectrophotometer (NanoDrop Technologies).

Real-Time PCR

Total RNA samples were subjected to reverse transcriptase reaction by using a GeneAmp RNA PCR kit (Applied Biosystems) according to the same method as used for semi-quantitative RT-PCR analysis. The expression levels of spliced XBP1 (method 1) and BiP mRNA were measured by real-time PCR using Platinum[®] SYBR[®] Green qPCR SuperMix UDG (Invitrogen Life Technologies, Carlsbad, CA). For measurement of spliced XBP1 in method 1, the double-stranded cDNA was synthesized from single-stranded cDNA by PCR in a total volume of 25 µL containing 10 pmol each of specific sense and anti-sense primers, 12.5 µL of Platinum[®] SYBR[®] Green qPCR SuperMix-UDG (2 x), and 0.5 µL of ROX Reference Dye. The PCR reaction was performed under the following thermal cycling conditions: 95°C 5 min, then 95°C 30 sec - 55°C 30 sec - 72°C 30 sec for 2 cycles. After treatment with *Pst*I for 1 hour as described for semi-quantitative RT-PCR, the gene-specific PCR products were measured continuously with an ABI PRISM 7900HT Sequence Detection System (Applied Biosystems) and quantitated based on standard curves obtained for the gene-specific cDNA cloned into pGEM[®]-T Easy vector (Promega Corporation). Data were adopted when the correlation coefficient

(*r* value) was more than 0.9. The PCR conditions were 95°C 5 min, then 95°C 30 sec - 55°C 30 sec - 72°C 30 sec for 40 cycles. For measurements of BiP mRNA with the real-time PCR system, gene-specific PCR products formed directly from single-stranded cDNA were measured. The PCR conditions were 95°C 5 min, then 95°C 30 sec - 55°C 30 sec - 72°C 30 sec for 40 cycles. The quantity of specific mRNA was normalized as a ratio to the amount of GAPDH mRNA. The values of fold increase over the control were calculated by use of the following formula: spliced XBP1 fold increase (% of control) = (normalized spliced XBP1 expression of chemical-treated cells/ normalized spliced XBP1 expression of vehicle control cells) × 100, and BiP fold increase (% of control) = (normalized BiP expression of chemical-treated cells/ normalized BiP expression of vehicle control cells) × 100.

Statistical analysis

The statistical significance of differences in spliced XBP1 and BiP mRNA expression between non-treated cells and chemical-treated cells were analyzed using the paired Student's *t* test (Snedecor and Cochran, 1989). The correlation coefficient (*r* value) was determined in this paper by the use of Pearson's correlation statistics (Snedecor and Cochran, 1989).

RESULTS AND DISCUSSION

The accumulation of unfolded or misfolded peptides activates three signaling pathways, mediated by IRE1, ATF6 and PERK. These UPR events can be followed by immunoblot analysis of ATF6 activation, phosphorylation of eIF2 α by PERK, and UPR-inducible proteins (such as BiP) and by RT-PCR analysis of XBP1 mRNA splicing mediated by IRE1 activation and UPR-inducible gene products. Shang reported quantitative methods for determination of these biomarkers based on immunoblotting, Northern blotting and RT-PCR (Shang, 2005). However, it is difficult to apply these methods to large numbers of samples. We can use quantitative real-time PCR analysis as a quantitative, simple and high-throughput assay, but it is difficult to measure activation of IRE1, ATF6 and PERK by real-time PCR analysis because the activations of these gene products involve phosphorylation or processing of protein. So we still require ER stress biomarkers that can be measured simply and quantitatively. Recently, the possibility has been suggested of widespread involvement of ER stress in neurodegenerative diseases, such as Alzheimer's and Parkinson's diseases

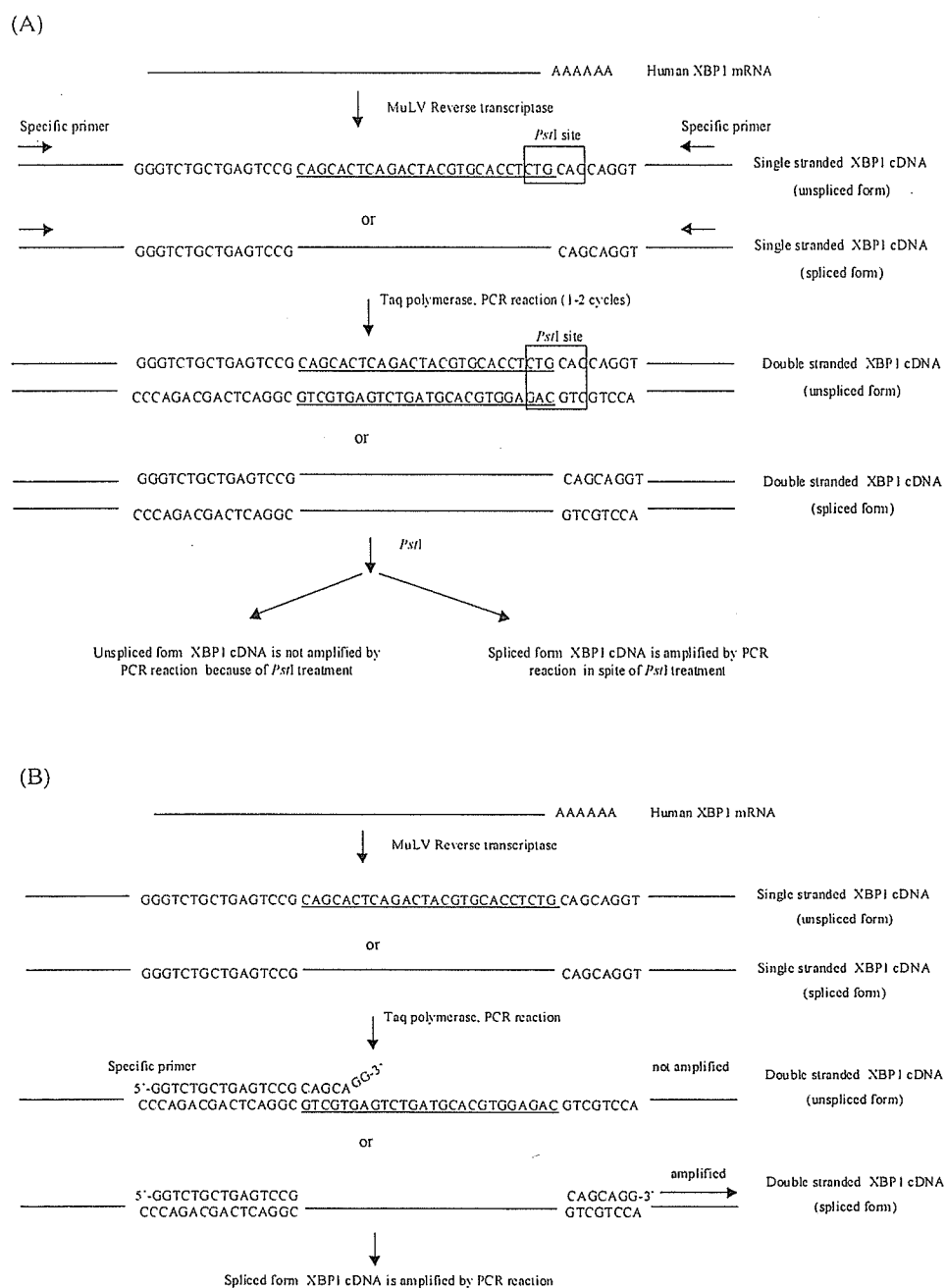


Fig. 1. The principle of spliced XBP1-specific RT-PCR ((A); method 1, (B); method 2). (A) Total RNA was reverse-transcribed and double-stranded cDNA was synthesized by PCR using specific sense and anti-sense primers for the XBP1 gene. The reaction mixture was treated with *Pst*I, then subjected to PCR reaction. The double-stranded cDNA derived from unspliced XBP1 mRNA was digested at the *Pst*I site, and so was not amplified by the PCR reaction. However, the double-stranded cDNA derived from spliced XBP1 mRNA was not digested, because of loss of the *Pst*I site owing to the splicing in response to ER stress, and was amplified by PCR. (B) Only the cDNA derived from spliced XBP1 mRNA was amplified by using a specific primer designed to span the 26 bp intron. The underlined nucleotides indicate the 26 bp intron sequence that is spliced out in response to ER stress. The sequence of XBP1 was taken from GeneBank (accession No. AB076383).

Quantitative measurement of spliced XBP1 mRNA.

(Kudo *et al.*, 2002; Ryu *et al.*, 2002; Schroder and Kaufman, 2005) and diabetes (Harding and Ron, 2002). Consequently quantitative methods to measure UPR events are needed for toxicological studies. Among biomarkers of ER stress, XBP1 gene expression is regulated by activated ATF6 protein and XBP1 protein is derived from spliced XBP1 mRNA. Further, spliced XBP1 mRNA is regulated by activation of IRE1. XBP1 protein derived from spliced XBP1 mRNA binds ERSE and UPR as a potent transcription factor. Therefore, we thought that it was more important to measure expression of spliced XBP1 as a biomarker that reflects the ATF6 signal and IRE1 signal than other biomarkers of ER stress. Until now, the measurement of spliced XBP1 expression has been mainly performed by semi-quantitative methods, such as agarose gel electrophoresis analysis after RT-PCR, because the difference in length between unspliced and spliced XBP1 is only 26 nucleotides (Yoshida *et al.*, 2001). In this study, we examined two methods (Fig. 1) for quantitative measurement of the expression of spliced XBP1 mRNA. As shown in Fig. 2 (A) (method 2), we could detect spliced XBP1 mRNA induction in

DTT- and Tm-treated THP-1 by using a specific primer designed to span the 26 bp intron. We could detect augmentation of spliced XBP1 more sensitively following *Pst*I treatment of double-stranded XBP1 cDNA (Fig. 2 (B), method 1). We thought that this detection was caused by *Pst*I treatment because we could not detect augmentation of spliced XBP1 under the condition of method 1 without *Pst*I treatment (Fig. 2 (A), total XBP1). As shown in Fig. 1 (A), the key points of this method are the synthesis of double-stranded cDNA and digestion by the restriction enzyme *Pst*I. Many restriction enzymes such as *Pst*I and *Eco*RI, can recognize specific double-stranded DNA sequences, but not single-stranded DNA. Although the recommended reaction conditions for *Pst*I include 100 mM salt concentration (Na^+ or K^+) in the reaction buffer (TaKaRa Bio on-line catalog), we used 50 mM K^+ in PCR in many case. It is reported that the enzyme activity of *Pst*I in buffer containing 50 mM sodium ion (Na^+) is 60% of that in buffer containing 100 mM buffer Na^+ , but the activity is higher in the presence of potassium ion (K^+) than Na^+ (TaKaRa Bio on-line catalog). Therefore, we concluded that *Pst*I was able to digest double-stranded

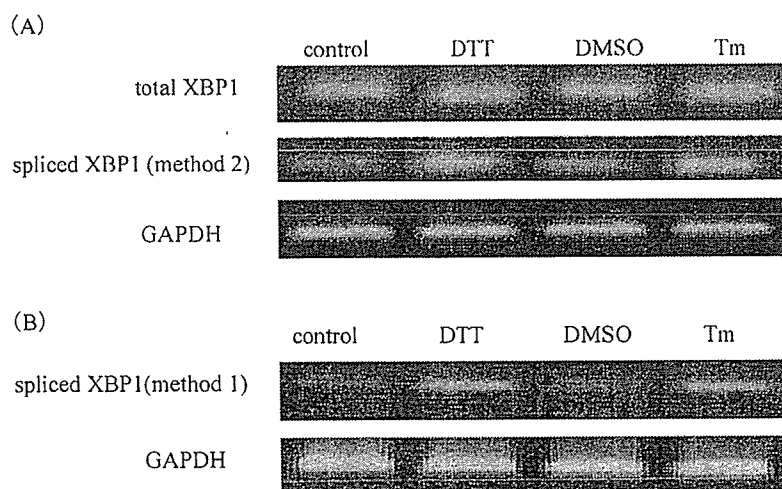


Fig. 2. Semi-quantitative RT-PCR analyses of spliced XBP1. (A) RT-PCR analyses of total XBP1 and spliced XBP1 mRNA were performed by using specific XBP1 primers (method 1 and method 2) without *Pst*I treatment. (B) RT-PCR analysis of spliced XBP1 (method 1) after *Pst*I treatment was performed as described in Materials and Methods. GAPDH mRNA expression was measured as an internal control under the same PCR conditions as used for XBP1 mRNA (methods 1 and 2). THP-1 cells (1×10^6 cells/mL) were exposed to dithiothreitol (DTT) 1 mM, tunicamycin (Tm) 3 $\mu\text{g}/\text{mL}$, or DMSO (solvent control) 2000 $\mu\text{g}/\text{mL}$ for 2 h. RNA was extracted, reverse-transcribed, and amplified by PCR with XBP1- and GAPDH-specific primers. PCR products were separated by electrophoresis on a 2.0% agarose gel.

cDNA under the general PCR buffer conditions.

We next investigated whether this method (method 1) could be applied to real-time PCR. We compared the amounts of spliced XBP1 cDNA determined by real-time PCR analysis with those determined by spectrophotometric analysis (Fig. 3). There was a linear correlation between the results of real-time PCR analysis and spectrophotometric analysis, and contaminating unspliced XBP cDNA seemed to have little influence. The divergence from the theoretical 1:1 correlation may be related to the differences in the measurement techniques.

Furthermore, we evaluated spliced XBP1 expression in THP-1 cells treated with representative ER stress agents, using a real-time PCR system. As shown in Fig. 4, we quantitatively detected induction of spliced XBP1 mRNA by ER stress agents, as well as expression of BiP mRNA. Shang and Lehrman found no correlation between the XBP1 splicing ratio and the levels of target transcripts such as BiP and EDEM (Shang and Lehrman, 2004). Here, we re-examined correlation between BiP gene expression and spliced XBP1 gene expression using the method developed in

this study. As shown in Fig. 5, we found a good correlation between the expression levels of spliced XBP1 and BiP gene mRNAs in THP-1 cells treated with ER stress agents ($r = 0.72$). We think that measurement of these mRNAs using the same platform is one of the reasons for the good correlation found here. In stressed ER cells, spliced XBP1 mRNA generated by the RNase activity of activated IRE1 is translated to a potent transcription factor that binds to ERSE and UPRE. Thus, our result is reasonable, because the promoter region of the BiP gene contains an ERSE region. Measuring the splicing ratio of XBP1 is expected to be a good biomarker for ER stress because of the good correlation with ATF6 (Shang and Lehrman, 2004). From our data and previous observations, it seems to be important to measure the level of spliced XBP1 mRNA to detect ER stress, as well as the splicing ratio of XBP1 mRNA.

Our method for measuring spliced XBP1 should lead to new insight into the involvement of ER stress in the pathogenesis of neurodegenerative diseases and diabetes.

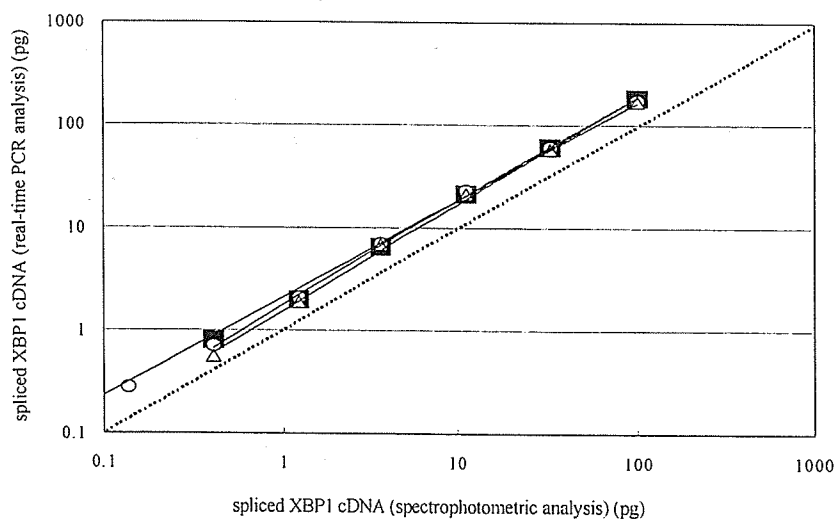


Fig. 3. Correlation between the amounts of spliced XBP1 cDNA determined by real-time PCR analysis and by spectrophotometric analysis. Spliced and unspliced XBP1 cDNAs cloned into pGEM[®]-T Easy vector (Promega Corporation) were used as template DNA. Real-time PCR analysis of spliced XBP1 cDNA was performed by method 1. The amounts of template spliced XBP1 cDNA and unspliced XBP1 cDNA were determined by measuring OD₂₆₀ on a NanoDrop spectrophotometer (NanoDrop Technologies). The amounts of unspliced XBP1 cDNA contaminated in spliced XBP1 template cDNA sample were 0 pg (Δ), 25 pg (\circ) and 50 pg (\blacksquare). Each value is the mean of two independent experiments. The dotted line represents the theoretical 1:1 correlation.

Quantitative measurement of spliced XBP1 mRNA.

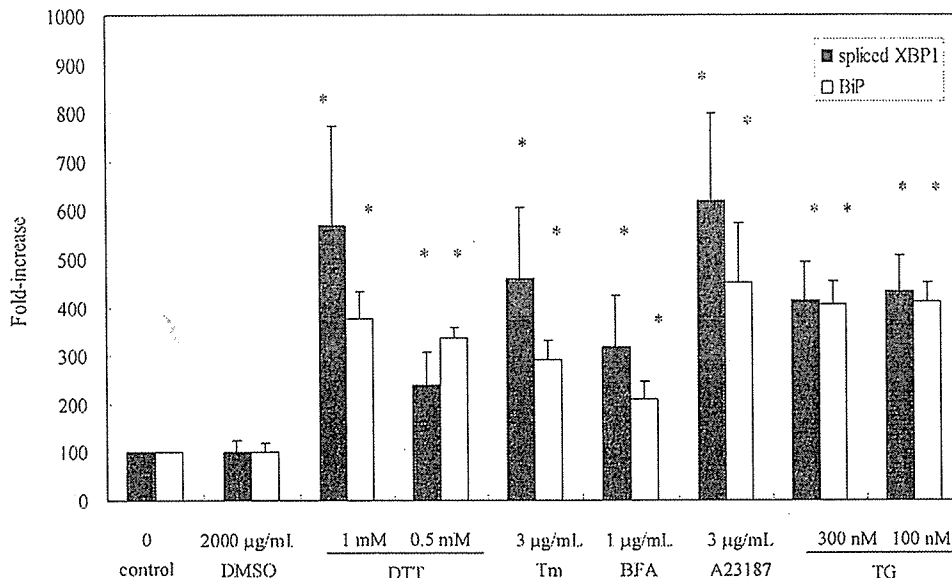


Fig. 4. Real-time PCR analysis of mRNA expression of spliced XBP1 (method 1) and BiP in THP-1 cells treated with ER stress agents. THP-1 cells (1×10^6 cells/mL) were exposed to ER stress agents for 2 hr. RNA was extracted, reverse-transcribed, and analyzed with the real-time PCR system after synthesis of double-stranded cDNA followed by *Pst*I treatment as described in Materials and Methods. The increases (fold) of mRNA levels were determined and normalized (Materials and Methods). Each value is the mean \pm S.D. of at least three independent experiments. Asterisks indicate a significant ($p < 0.05$) difference between chemical-treated cells and vehicle-treated cells.

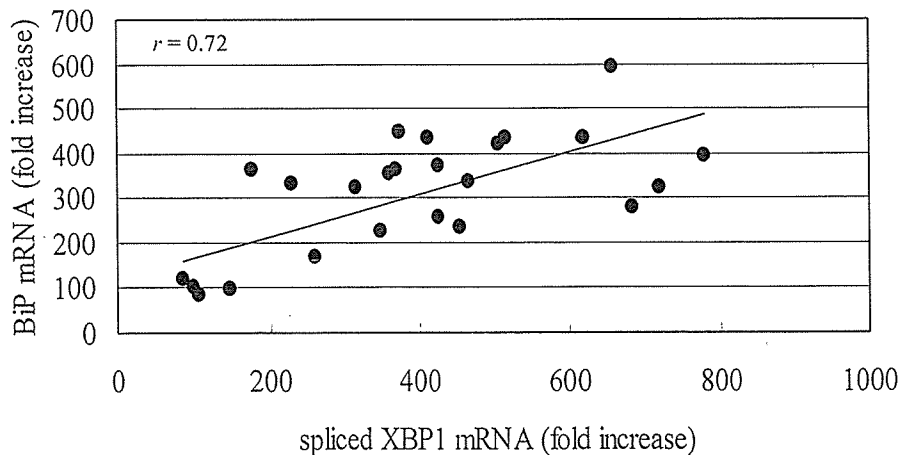


Fig. 5. Correlation between expression levels of spliced XBP1 and BiP mRNAs in THP-1 cells treated with ER stress agents. The increases (fold) of mRNA levels were determined and normalized as described in Materials and Methods. Each data point represents the result of a single experiment. The correlation coefficient (r value) was 0.72.

REFERENCES

- Ellgaard, L. and Helenius, A. (2003): Quality control in the endoplasmic reticulum. *Nat. Rev. Mol. Cell Biol.*, **4**, 181-191.
- Gething, M.J. (1999): Role and regulation of the ER chaperone BiP. *Cell Dev. Biol.*, **10**, 465-472.
- Harding, H.P. and Ron, D. (2002): Endoplasmic reticulum stress and the development of diabetes: a review. *Diabetes*, **51**, S455-461.
- Harding, H.P., Zhang, Y. and Ron, D. (1999): Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. *Nature*, **397**, 271-274.
- Haze, K., Yoshida, H., Yanagi, H., Yura, T. and Mori, K. (1999): Mammalian transcription factor ATF6 is synthesized as a transmembrane protein and activated by proteolysis in response to endoplasmic reticulum stress. *Mol. Biol. Cell.*, **10**, 3787-3799.
- Kudo, T., Katayama, T., Imaizumi, K., Yasuda, Y., Yatera, M., Okochi, M., Tohyama, M. and Takeda, M. (2002): The unfolded protein response is involved in the pathology of Alzheimer's disease. *Ann. N. Y. Acad. Sci.*, **977**, 349-355.
- Liu, C.Y., Schroder, M. and Kaufman, R.J. (2000): Ligand-independent dimerization activates the stress response kinases IRE1 and PERK in the lumen of the endoplasmic reticulum. *J. Biol. Chem.*, **275**, 24881-24885.
- Mori, K., Ma, W., Gething, M.J. and Sambrook, J. (1993): A transmembrane protein with a cdc2+/CDC28-related kinase activity is required for signaling from the ER to the nucleus. *Cell*, **74**, 743-756.
- Ryu, E.J., Harding, H.P., Angelastro, J.M., Vitolo, O.V., Ron, D. and Greene, L.A. (2002): Endoplasmic reticulum stress and the unfolded protein response in cellular models of Parkinson's disease. *J. Neurosci.*, **22**, 10690-10698.
- Schroder, M. and Kaufman, R.J. (2005): ER stress and the unfolded protein response. *Mutat. Res.*, **569**, 29-63.
- Shamu, C.E. and Walter, P. (1996): Oligomerization and phosphorylation of the Ire1p kinase during intracellular signaling from the endoplasmic reticulum to the nucleus. *EMBO J.*, **15**, 3028-3039.
- Shang, J. (2005): Quantitative measurement of events in the mammalian unfolded protein response. *Methods*, **35**, 390-394.
- Shang, J. and Lehrman, M.A. (2004): Discordance of UPR signaling by ATF6 and Ire1p-XBP1 with levels of target transcripts. *Biochem. Biophys. Res. Commun.*, **317**, 390-396.
- Shen, J., Chen, X., Hendershot, L. and Prywes, R. (2002): ER stress regulation of ATF6 localization by dissociation of BiP/GRP78 binding and unmasking of Golgi localization signals. *Dev. Cell.*, **3**, 99-111.
- Shi, Y., Vattem, K.M., Sood, R., An, J., Liang, J., Stramm, L. and Wek, R.C. (1998): Identification and characterization of pancreatic eukaryotic initiation factor 2 alpha-subunit kinase, PEK, involved in translational control. *Mol. Biol. Cell.*, **18**, 7499-7509.
- Sidrauski, C., Chapman, R. and Walter, P. (1998): The unfolded protein response: an intracellular signalling pathway with many surprising features. *Trends Cell Biol.*, **8**, 245-249.
- Snedecor, G.M. and Cochran, W.G. (1989): *Statistical Methods*. 8th edn. Ames, IA, Iowa State University Press.
- TaKaRa Bio on-line catalog, <http://www.takara-bio.co.jp/>
- Welihinda, A.A. and Kaufman, R.J. (1996): The unfolded protein response pathway in *Saccharomyces cerevisiae*. Oligomerization and transphosphorylation of Ire1p (Ern1p) are required for kinase activation. *J. Biol. Chem.*, **271**, 18181-18187.
- Yoshida, H., Haze, K., Yanagi, H., Yura, T. and Mori, K. (1998): Identification of the cis-acting endoplasmic reticulum stress response element responsible for transcriptional induction of mammalian glucose-regulated proteins. Involvement of basic leucine zipper transcription factors. *J. Biol. Chem.*, **273**, 33741-33749.
- Yoshida, H., Matsui, T., Yamamoto, A., Okada, T. and Mori, K. (2001): XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. *Cell*, **107**, 881-891.
- Zhang, K. and Kaufman, R.J. (2004): Signaling the unfolded protein response from the endoplasmic reticulum. *J. Biol. Chem.*, **279**, 25935-25938.



Development of an in vitro skin sensitization test using human cell lines: The human Cell Line Activation Test (h-CLAT)

I. Optimization of the h-CLAT protocol

T. Ashikaga^{a,*}, Y. Yoshida^b, M. Hirota^a, K. Yoneyama^a, H. Itagaki^a, H. Sakaguchi^b,
M. Miyazawa^b, Y. Ito^b, H. Suzuki^b, H. Toyoda^a

^a Safety and Analytical Research Center, Shiseido Co., Ltd., 2-12-1 Fukuura, Kanazawa-ku, Yokohama-shi, Kanagawa 236-8643, Japan

^b Safety and Microbial Control Research Center, Kao Corporation, 2606 Akabane Ichikai-Maich, Haga-Gun, Tochigi 321-3497, Japan

Received 20 April 2005; accepted 17 October 2005

Available online 28 November 2005

Abstract

The aim of this study is to optimize the experimental conditions for an in vitro skin sensitization test using the human cell lines THP-1 and U-937. As regards pre-culturing time, the expression of CD86 on DNCB-treated THP-1 cells tended to be higher after 48 h and 72 h pre-culture compared with other time points evaluated. Next, we investigated the effect of chemical treatment time, and found that induction of CD86 expression on THP-1 cells by DNCB reached a plateau after 24 h. Augmentation of CD86 expression is often observed when cells are treated with a subtoxic dose of allergens. To determine the appropriate dose of test samples, the cytotoxicity of test samples to THP-1 and U-937 cells was assessed with MTT assay, and the 50% inhibitory concentration (IC₅₀) of each test sample was calculated. Based on the cytotoxicity assay data, four concentrations in the range between toxic and non-toxic were selected (0.1×, 0.5×, 1× and 2× IC₅₀). Several kinds of antibodies were tested for staining THP-1 and U-937 cells treated with allergens/non-allergens (e.g., DNCB, Ni/SLS), and suitable antibodies for staining CD86 and CD54 were selected. We confirmed that the working dilutions of the selected CD86 and CD54 antibodies were appropriate for use in our method. The effect of an FcR blocking procedure was also evaluated. The mean fluorescence intensity (MFI value) was decreased by the FcR blocking procedure, which indicated that non-specific staining was blocked. Therefore, this procedure should be included in the method. Based on our findings, the protocol for this assay was optimized and the experimental conditions to be used in a future validation study were identified. We propose to call this kind of in vitro skin sensitization test h-CLAT, which is short for human Cell Line Activation Test.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: CD86; CD54; Alternative methods; Cell line; In vitro skin sensitization; THP-1; U-937

Abbreviations: ATCC, American type culture collection; APC, antigen presenting cell; BSA, bovine serum albumin; DCs, dendritic cells; DMSO, dimethyl sulfoxide; DNCB, dinitrochlorobenzene; FBS, fetal bovine serum; FcR, Fc receptor; IC₅₀, 50% inhibitory concentration; LCs, Langerhans cells; mAb, monoclonal antibody; MFI, mean fluorescence intensity; MTT, methylthiazolyl-diphenyl-tetrazolium bromide; Ni, nickel sulfate hexahydrate; PBMC, peripheral blood mononuclear cell; PI, propidium iodide; *p*PD, *p*-phenylenediamine; RFI, relative fluorescence intensity; SLS, sodium lauryl sulfate.

* Corresponding author. Tel.: +81 45 788 7309/7308; fax: +81 45 788 7309/7295.

E-mail address: takao.ashikaga@to.shiseido.co.jp (T. Ashikaga).

1. Introduction

There have been a number of attempts to develop non-animal alternative methods for skin sensitization testing (De Silva et al., 1996). Examination of the phenotypic and functional changes induced in antigen-presenting cells (APC), including Langerhans cells (LCs), by test agents is one of the most important approaches for developing in vitro sensitization methods. However, it is difficult to isolate LCs from skin. In recent years, the development

of culture techniques to generate dendritic cells (DCs) from human peripheral blood (PBMC-derived DC) has provided a source of LC-like APCs. Therefore, PBMC-derived DCs have been used for in vitro sensitization assays as professional antigen-presenting cells. These assays are based on the phenotypic changes in DCs, such as increase of CD86 or CD54 expression (Aiba et al., 1997), IL-1 β release (Hopper et al., 1995) and internalization of MHC class II molecules (Becker et al., 1997). However, obtaining DCs from peripheral blood not only takes several days, but also requires complicated procedures. Furthermore, the effects of chemicals on the surface phenotype of DCs were dependent on the source of peripheral blood, i.e., they varied from donor to donor (Aiba et al., 1997; Hulette et al., 2002). Such drawbacks would hamper efforts to validate such a method and limit its use as an alternative to animal testing. In our previous study, we found that CD86 and/or CD54 expression of THP-1 cells (monocytic leukemia cell line) or U-937 (histiocytic lymphoma cell line) was enhanced after a 24 h exposure to allergens, but not to non-allergens (Ashikaga et al., 2002, 2003; Yoshida et al., 2003). These cell lines are easy to culture and the cells retain well various key in vivo characteristics, such as esterase activity, lysozyme production and phagocytosis (Tsuchiya et al., 1980; Sundstrom and Nilsson, 1976). We considered that such human cell lines could be used as replacements for DCs in an in vitro skin sensitization test. There have been some similar attempts using other cell lines such as KG-1 (acute myelogenous leukemia cell line) (Hulette et al., 2002) and XS52 (Langerhans-like murine cell line) (Neisuis et al., 1999). However, a lot of work remains to be done in order to make the cell line system suitable for practical use as an alternative to animal tests. In this study, we examined various test conditions in order to optimize the cell line activation test using THP-1 or U-937 and thereby to develop a suitable protocol for use in a future validation study.

2. Materials and methods

2.1. Cells and culture

Both THP-1 and U-937 cells were purchased from American Type Culture Collection (ATCC, Manassas, VA, USA). Cells were cultured in RPMI 1640 medium (Invitrogen Corp., Carlsbad, CA, USA) with 10% FBS (v/v) (JRH Biosciences, Lenexa, KS, USA), 0.05 mM 2-mercaptoethanol and 1% antibiotic-antimycotic mixture (Invitrogen Corp., Carlsbad, CA, USA).

2.2. MTT assay

The MTT assay was used to evaluate the cytotoxicity of nine test chemicals. Dinitrochlorobenzene (DNCB), *p*-phenylenediamine (*p*PD), 2-mercaptobenzothiazole (2-MBT), nickel sulfate hexahydrate (Ni), cobalt sulfate heptahydrate (Co), ammonium tetrachloroplatinate (Pt),

sodium lauryl sulfate (SLS), Tween 80 and dimethyl sulfoxide (DMSO) were purchased from Sigma-Aldrich. *p*PD, Ni, Co, Pt, SLS, and Tween 80 were first solubilized in saline, while DNCB and 2-MBT were solubilized in DMSO. Cells were seeded at 1×10^5 cells/100 μ L/well in a 96-well plate and cultured with test chemicals for 24 h at various concentrations. A 25 μ L aliquot of MTT solution (1.67 mg/mL, dissolved in the medium) was applied to each well. After reaction for 3 h at 37 $^{\circ}$ C, the absorbance at both 570 nm and 660 nm was measured. The IC₅₀ values of the test chemicals were as follows: DNCB 5 μ g/mL, *p*PD 90 μ g/mL, 2-MBT 170 μ g/mL, Ni 170 μ g/mL, Co 130 μ g/mL, Pt 100 μ g/mL, SLS 90 μ g/mL, Tween 80 1800 μ g/mL; DMSO greater than 5000 μ g/mL. From the IC₅₀ values, four concentrations, i.e., 0.1 \times , 0.5 \times , 1 \times , and 2 \times IC₅₀, were selected as test doses. If a test chemical has deoxidizing activity, it will be necessary to wash the cells before addition of MTT solution to avoid direct interaction of the test chemical with MTT.

2.3. Pre-culture

Cells were seeded at 0.4×10^6 cells/mL and cultured for 24 h, 48 h, 72 h or 96 h in a cell culture flask. After the incubation, cells were collected, and 1×10^6 cells/mL were exposed to 1 μ g/mL or 3 μ g/mL DNCB for 24 h in fresh medium.

2.4. Chemical treatment

Cells were seeded at 1×10^6 cells/mL or 0.5×10^6 cells/mL in a 24-well plate and cultured with chemicals for 8 h, 12 h, 24 h, 48 h or 72 h. When DMSO was used as a solvent, its final concentration in culture media was less than 0.2%.

2.5. Flow cytometric analysis and antibodies

Cell staining was done at 4 $^{\circ}$ C for 30 min, using several kinds of commercially available monoclonal antibodies. The following two were chiefly used: anti-human CD86 antibody from BD-PharMingen (clone Fun-1, San Diego, CA, USA) and anti-human CD54 antibody from DAKO (clone 6.5B5, Glostrup, Denmark). FITC labeled-mouse IgG1 was purchased from DAKO (clone DAK-G01, Glostrup, Denmark) and used as an isotype control. Chemical-treated cells were washed twice with PBS containing 0.1% BSA. Then, cells were treated with 0.01% globulins, Cohn fraction II, III (SIGMA-Aldrich) for 10 min on ice to block FcR. The cells were washed once with PBS containing 0.1% BSA, and the expression of cell surface antigens was analyzed by flow cytometry. Dead cells were gated out by staining with propidium iodide (PI, 0.625 μ g/ml). Flow cytometric analysis was performed with a FACSCalibur CellQuest. (Becton Dickinson, San Jose, CA, USA.) at Kao or an EPICS XL-MCL System II, (Beckman Coulter Co., Ltd. Fullerton, CA, USA) at Shise-

ido. In total, 10,000 living cells were analyzed. When the cell viability was less than 50%, the relative fluorescence intensity (RFI) was not calculated, because of diffuse labeling of cytoplasmic structures as a result of cell membrane destruction (Becker et al., 1994). RFI was used as an indicator of CD86 and CD54 expression, and was calculated as follows:

$$\frac{\text{MFI of chemical-treated cells} - \text{MFI of chemical-treated isotype control cells}}{\text{MFI of vehicle control cells} - \text{MFI of vehicle isotype control cells}}$$

MFI = (geometric) mean fluorescence intensity.

3. Results

3.1. Effect of pre-culture time on CD86 expression

The effect of pre-culture time on the CD86 expression of THP-1 cells was studied (Fig. 1). DNCB (3 $\mu\text{g}/\text{mL}$), used as a positive control, enhanced CD86 expression two- to three-fold over the control independently of the pre-culture time. However, the expression of CD86 tended to be higher at 48 h and 72 h pre-culture compared to other time points evaluated, so 48 h or 72 h was considered a suitable pre-culture time. No augmentation of CD86 expression was observed with a lower concentration of DNCB (1 $\mu\text{g}/\text{mL}$).

3.2. Effect of chemical treatment time on CD86 expression

We investigated the optimum exposure time of cells to allergens for maximum induction of CD86 expression. After pre-culture, THP-1 cells seeded at 1×10^6 cells/mL were treated with 1 $\mu\text{g}/\text{mL}$ or 3 $\mu\text{g}/\text{mL}$ DNCB for 6 h, 12 h, 24 h, 48 h or 72 h. As shown in Fig. 2, induction of CD86 expression on THP-1 cells increased with treatment time, reaching a plateau at 24 h in the case of DNCB

(3 $\mu\text{g}/\text{mL}$) treatment. DNCB at 1 $\mu\text{g}/\text{mL}$ did not cause any change of CD86 expression even after 72 h.

3.3. Initial cell concentration

THP-1 cells were seeded at 1×10^6 cells/mL or 0.5×10^6 cells/mL in 1 mL of medium in a 24-well plate

and treated with two concentrations of DNCB or Ni (allergens) or SLS (a non-allergen). After exposure to these chemicals for 24 h or 48 h, expression of CD86 and CD54 on the surface of the THP-1 cells was measured (Fig. 3). In some cases (arrowed in Fig. 3), 1×10^6 cells/mL was better as an initial cell concentration. Indeed, in the case of CD54 expression on cells treated with Ni for 24 h, augmentation was observed only at 1×10^6 cells/mL. In the case of U-937 cells, there seemed to be no difference between the results with 1×10^6 cells/mL and 0.5×10^6 cells/mL (data not shown).

3.4. Dose setting based on the results of cytotoxicity test

We have previously observed augmentation of CD86 expression when cells were treated with a subtoxic dose of allergens. Therefore, to determine appropriate test concentrations, the 50% inhibitory concentration (IC₅₀) of each test sample was calculated by means of MTT assay. The survival rates at $0.1 \times \text{IC}_{50}$, $0.5 \times \text{IC}_{50}$ and $2 \times \text{IC}_{50}$ were calculated from the growth inhibition curves of the chemicals. As shown in Table 1, four concentrations ($0.1 \times$, $0.5 \times$, $1 \times$ and $2 \times$ of the IC₅₀) adequately covered the range from toxic (under 50% cell viability) to non-toxic concentrations. Therefore, we chose these four concentrations

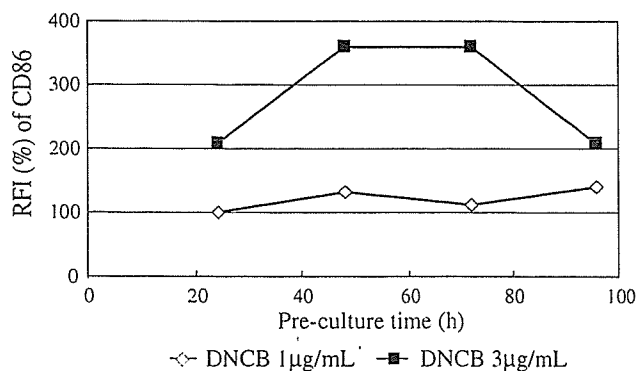


Fig. 1. Effect of pre-culture time on CD86 expression. THP-1 cells were seeded at 0.4×10^6 cells/mL and cultured for 24 h, 48 h, 72 h or 96 h. After the incubation, cells were transferred to fresh medium, and exposed to DNCB for 24 h. RFI values are mean values of two experiments.

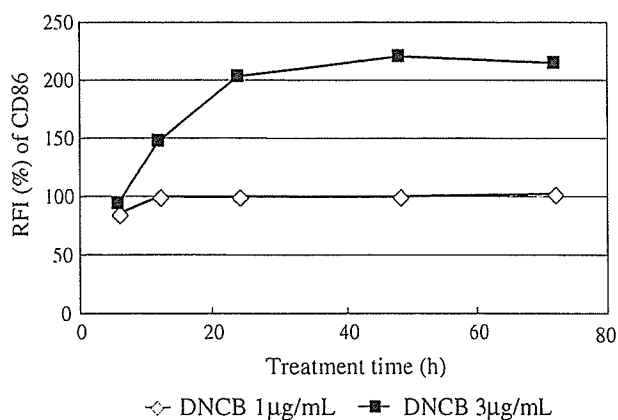


Fig. 2. Effect of chemical treatment time on CD86 expression. THP-1 cells were treated with DNCB for 6 h, 12 h, 24 h, or 72 h. After exposure to DNCB, expression of CD86 was measured. RFI values are mean values of two experiments.

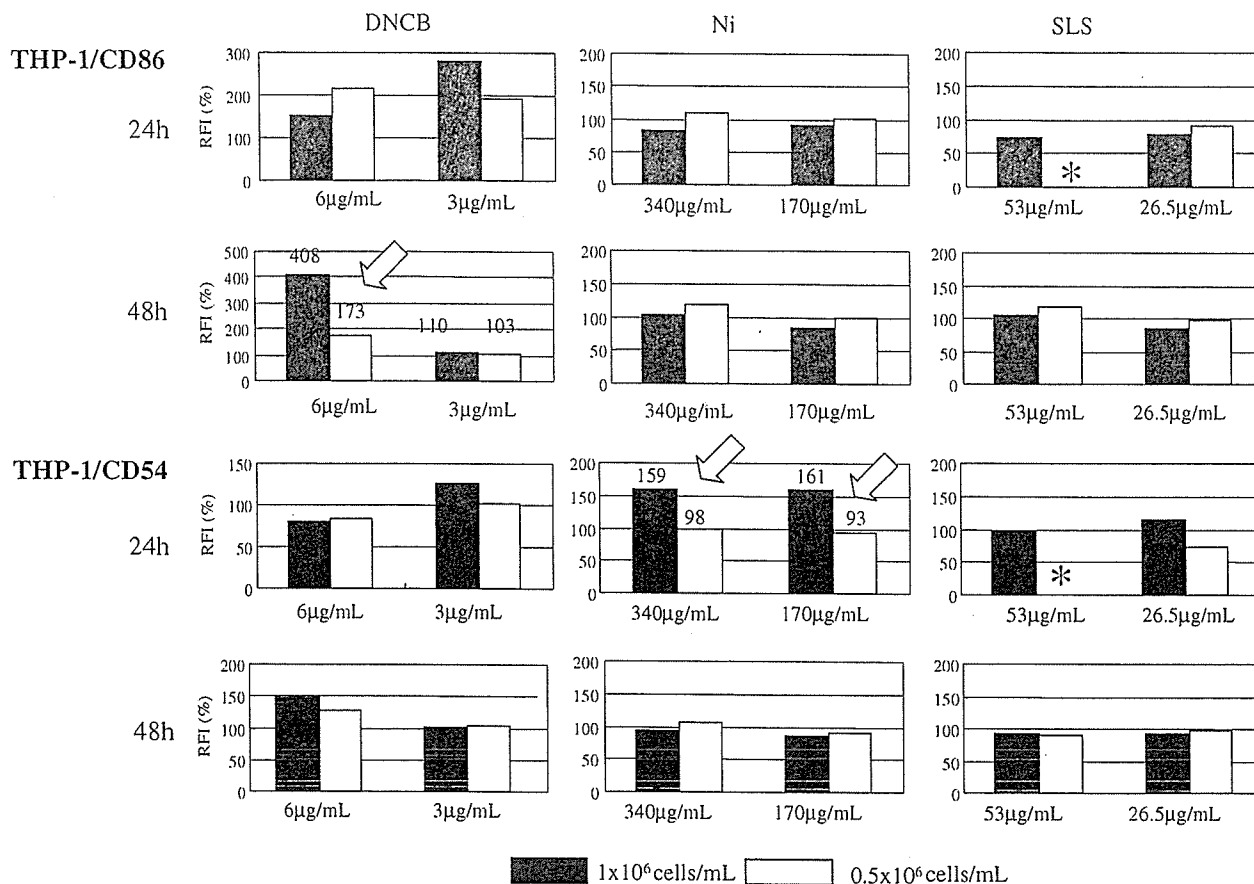


Fig. 3. Initial number of cells seeded. THP-1 cells were seeded at 1×10^6 cells/mL or 0.5×10^6 cells/mL and treated with DNCB, Ni or SLS for 24 h or 48 h. CD86 and CD54 expression levels at the two initial cell concentrations were compared. The arrows in figure indicate large differences. RFI values are means of two experiments. *; not determined.

Table 1
Dose setting based on the results of cytotoxicity testing

Calculated cell viability (%)	THP-1 (24 h/48 h)			U-937 (24 h/48 h)		
	0.1× MC50	0.5× IC50	2× IC50	0.1× MC50	0.5× IC50	2× IC50
>90	8/8	2/2	0/0	7/5	3/4	0/0
80–90	0/0	1/0	0/0	1/1	0/0	0/0
70–80	0/0	5/4	0/0	0/0	5/2	0/0
60–70	0/0	0/2	0/0	0/2	0/1	0/0
50–60	0/0	0/0	0/0	0/0	0/1	0/0
40–50	0/0	0/0	0/0	0/0	0/0	0/0
30–40	0/0	0/0	3/1	0/0	0/0	1/0
<30	0/0	0/0	5/7	0/0	0/0	7/8

Cytotoxicity of eight chemicals (except DMSO) was investigated by MTT assay and the values of 50% inhibitory concentration (IC50) were calculated. The survival rate at 2× IC50, 0.5× IC50 or 0.1× IC50 was calculated from the growth inhibition curves of these eight chemicals.

as application doses for evaluation of induction of CD86 and CD54 expression. The IC50 of DMSO was estimated to be over 2500 μ g/mL, which was the highest concentration under our test conditions.

3.5. Selectivity of antibody

The reactivity of various kinds of anti-CD86 and anti-CD54 antibodies was examined. THP-1 cells were treated with DNCB, Ni or pPD (allergens) or SLS (non-allergen)

for 24 h, then were stained with each antibody at the recommended concentration. As shown in Fig. 4, antibody D exhibited the highest reactivity for staining CD 86 in the allergen-treated cells, among the antibodies we tested. For staining CD54, antibody G showed the highest reactivity, except in the case of DNCB-treated cells. Based on these findings, we chose antibody D (BD-PharMingen, clone Fun-1) for CD86 staining and antibody G (DAKO, clone 6.5B5) for CD54 staining in the following experiments.

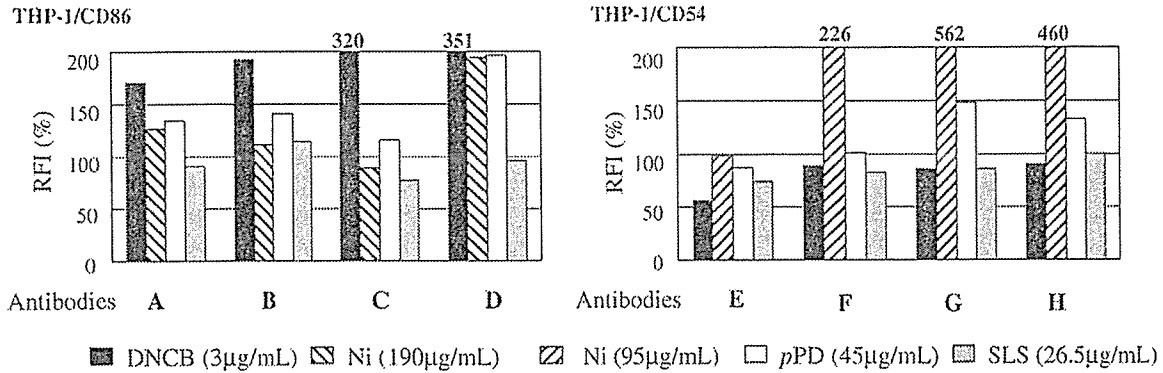


Fig. 4. Antibody dependence. THP-1 cells were seeded at 1×10^6 cells/mL, treated with each chemical for 24 h, and stained with the antibodies. The RFI values were calculated and compared.

3.6. Determination of the optimal amounts of antibodies

We tested whether the recommended dilutions of the selected antibodies were suitable for this assay or not. THP-1 cells were treated with DNCB (positive control for CD86), pPD (positive control for CD54) or SLS (negative control for both CD86 and CD54) for 24 h, then stained with antibodies D and G at three concentrations (3×, 1× and 1/3× recommended dilution). As shown in Fig. 5, the results showed no significant dependence on antibody dilution. This result suggested that the working dilutions recommended in the data sheets of antibodies D and G were appropriate for use in our assay. The recommended dilutions were used in the following experiments.

3.7. Effect of FcR-blocking on cell-surface antigen expression of THP-1

THP-1 cells were seeded at 1×10^6 cells/mL in a 24-well plate and cultured for 24 h without any chemical treatment. The cells were washed, then FcR-blocking was done or not done, and the cells were stained with anti-CD86 antibodies (A, B or D) or anti-CD54 antibodies (E or G). The results are shown in Fig. 6. For all antibodies exam-

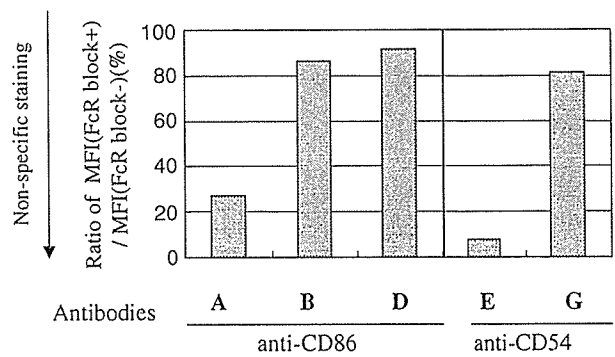


Fig. 6. Effect of FcR-blocking on measurement of THP-1 cell surface antigens. THP-1 cells were cultured without chemicals. Before staining of CD86 or CD54, cells were treated with or without 0.01% human globulins to block non-specific binding. Then, cells of both groups were stained with various antibodies. The ratio of MFI (FcR block+) and MFI (FcR-), which is a measure of the specificity of the antibody, was calculated.

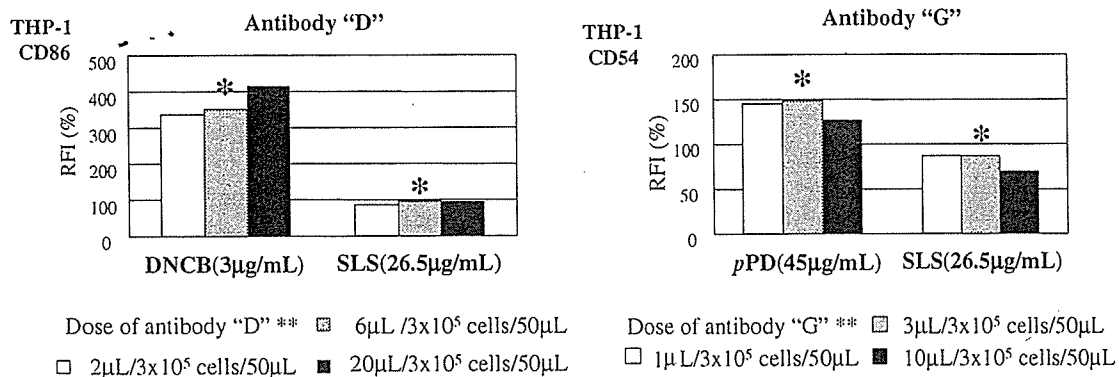


Fig. 5. Determination of the optimal amount of antibodies for application DNCB or SLS-treated THP-1 cells were stained with both antibodies (D and G) at various concentrations. *; recommended dilution (see methods). **; volume of antibody/number of cells/total volume.