The Embryonic Stem Cell Test for the Early Selection of Pharmaceutical Compounds

Sonia Whitlow, Heinrich Bürgin and Nicole Clemann
F. Hoffmann-La Roche Ltd., Pharma Research, Basel, Switzerland

Summary
Potential teratogenicity is a major consideration in the development of pharmaceutical substances. Currently its assessment involves large numbers of animal tests at high cost. This study assessed the feasibility of using the embryonic stem cell test (EST), validated by ECVAM in 1999, as a tool for the prediction of embryonic toxicity of pharmaceutical substances early in their development programmes. ESTs were carried out on 6 chemicals with well established toxicity characteristics established from literature and from the ECVAM study, and then on 10 Roche internal pharmaceutical substances already tested in vivo. The model correctly classified 81% of the substances. Further experiments are necessary to increase the database of the assay.

Zusammenfassung: Der embryonale Stammzelltest für die Auswahl von pharmazeutischen Substanzen
Während der Entwicklung pharmazeutischer Substanzen spielt ihre potentielle Teratogenität eine wichtige Rolle. Zurzeit erfordert die Überprüfung der Teratogenität eine grosse Anzahl von Tierversuchen, die mit hohen Kosten verbunden sind. Das Ziel dieser Studie war es, das Potential des von ECVAM im Jahr 1999 validierten embryonalen Stammzelltests (EST) zur Vorhersage des teratogenen Potentials von pharmazeutischen Substanzen in deren frühen Entwicklung zu ermitteln. Der EST wurde mit sechs Chemikalien durchgeführt, die bereits in der ECVAM Studie getestet wurden und mit zehn pharmazeutischen Substanzen von Roche, die auch schon in vivo getestet wurden. Das Modell klassifizierte 81% der Substanzen richtig. Zur Vergrösserung der Datenbasis des EST sind weitere Experimente notwendig.

1 Introduction

It has been known for more than 20 years that embryonic stem cells can differentiate *in vitro* into various tissues of the endoderm, ectoderm and mesoderm. These cells can be maintained indefinitely in an undifferentiated state in the presence of leukaemia inhibiting factor (LIF, Williams et al., 1988), and can be frozen and thawed without losing their pluripotent character. Upon removal of LIF from the culture, the cells grow and differentiate into various tissues, including contracting cardiomyocytes (Doetschman et al., 1985; Rohwedel et al., 2001; Wobus et al., 1991, 2005). The latter cells are easily visible under the microscope so that effects on their development caused by the addition of test substances to the culture medium can be readily investigated. The degree of inhibition of the development of contracting cardiomyocytes could be used as a measure of potential embryotoxicity, and thus the potential of the test substance to cause malformations, in other words, it could be used to identify a potentially teratogenic substance.

Embryonic stem cell tests using mouse cells (ESTs) were developed independently by Newall et al. (1994) and later by Spielmann et al. (1997). They investigated the effects of various chemicals on the viability of embryonic stem (ES) cells and differentiated adult fibroblast cells, as in an earlier study by Laschinsky et al. (1991), as well as their effects on the ability of ES cells to differentiate into contracting cardiomyocytes. Three endpoints were determined: 50% inhibition of growth of both ES cells and fibroblasts, and 50% inhibition of differentiation into cardiomyocytes. The EST continued to be developed further and in 1997, Spielmann et al. investigated 16 chemicals with known *in vivo* embryotoxic potential. As part of the investigation, biostatistical models were developed to predict the *in vivo* embryotoxic potential of the test chemicals by classification into 3 different classes: non-, weakly, and strongly embryotoxic. The predictions of embryotoxicity obtained from the ESTs were compared to the *in vivo* results and agreed in all cases.

Since then the EST has been refined (Seiler et al., 2004, 2006) and prediction models have been further developed to estimate toxic levels of chemicals *in*

vivo. The EST, along with the whole embryo culture assay and the limb bud micromass assay, have been validated in an extensive study sponsored by ECVAM, the European Centre for the Validation of Alternative Methods (Scholz et al., 1999b; Genschow et al., 2000, 2002 and 2004). The EST assays resulted in an initial set of 6 chemicals being 93% correctly classified, and 78% of the 14 chemicals were correctly classified in the formal validation study.

This paper describes the use of the EST for pharmaceutical substances by studying 16 compounds, of which 10 were pharmaceutical substances developed by F. Hoffmann-La Roche AG and 6 had been tested already in the ECVAM validation study (Genschow et al., 2000, 2002, 2004). Although only a limited number of compounds was investigated, the accuracy of the model predictions was sufficiently encouraging to warrant further development of the model to improve its database. An accurate and reliable prediction model would allow the screening of substances to give a ranking of the degree of embryotoxicity early in compound development. The EST could help to reduce animal experiments by replacing the *in vivo* screening model used in some laboratories.

2 Materials and methods

Test chemicals
Six chemicals were selected from the EC-VAM list on the basis of their embryotoxicity and reliable results from *in vivo* investigations. They included two strongly (5-bromo-2-deoxyuridine, *all-trans*-retinoic acid), two weakly (lithium chloride, boric acid) and two non-embryotoxic (penicillin G sodium salt, saccharin sodium hydrate) chemicals. The 10 pharmaceutical substances tested were developed by F. Hoffmann-La Roche AG and had also been tested extensively *in vivo*. The compounds were derived from different chemical classes (e.g. piperidine, pyridine, cyclic peroxide type) and were developed against different targets. Of these, five were strongly embryotoxic:
R1 – an antimalarial substance which caused higher levels of intrauterine deaths in rats and rabbits, and blood

vessel and bone development anomalies in the rat.
R2 – an analgesic which induced oral clefts and an increased rate of intrauterine deaths.
R3 – a collagenase inhibitor which caused an increased incidence of intrauterine deaths, and severe craniofacial, cardiovascular and urogenital malformations.
R4 – an oxidosqualene cyclase inhibitor which is embryolethal, causing a high level of intrauterine death and incidences of cleft palate.
R5 – an oxidosqualene cyclase inhibitor causing malformations including cleft palate.

Four substances were classified as being non-embryotoxic:
R6 – an inhibitor of monamine oxidase B.
R7 – an oral cephalosporin, which was maternally toxic, but not embryotoxic.
R8 – an analgesic and an enantiomer of R2.
R9 – an ACE-inhibitor.

The remaining substance was weakly embryotoxic; it was the only one that had been so classified in the *in vivo* experiments:
R10 – a calcium antagonist which caused cardiovascular anomalies.

Cell culture
The two types of cells, Balb/c 3T3 cells (adult mouse fibroblast cells), clone A31, and embryonic stem cells (mouse), D3, were obtained from the American Type Culture Collection (ATCC), and were cultured by the method developed by Williams et al. (1988). The fibroblast medium was DMEM (Gibco) supplemented with 10% FCS (Sigma), 4 mM glutamine, 50 U/ml penicillin, and 50 μg/ml streptomycin (the latter three from Gibco). ES cells were cultured by the same method, but using DMEM supplemented with 20% FCS, 2 mM glutamine, 50 U/ml penicillin, 50 μg/ml streptomycin, 1% non-essential amino acids (all from Gibco), 0.1 mM β-mercaptoethanol (Sigma), and 1000 U/ml leukaemia inhibitory factor (Sigma), which was added directly to the plates. All cells were detached using 0.5% trypsin/EDTA (Gib-

co), counted and resuspended in culture medium three times per week.

Assays
The assay methods were those developed by Spielmann et al. 1997 and updated in INVITTOX, protocol no. 113 (1999). All tests on the chemicals were carried out "blind".

Biostatistical analysis and prediction model
Data files of optical densities (OD_{570}) generated by a microplate reader were copied into an EXCEL spreadsheet. Mean OD values, standard deviations and viabilities were calculated automatically. The following assay endpoints were calculated graphically from the concentration-response curve in the spreadsheet:
$IC_{50}D3$ – the concentration of test substance at which 50% of D3 cells died
$IC_{50}3T3$ – the concentration of test substance at which 50% of the 3T3 cells died.
$ID_{50}D3$ – the concentration of test substance at which there was a 50% reduction in the differentiation of D3 cells into contracting cardiomyocytes.

The IC_{50} values of the D3 and 3T3 cells from the cytotoxicity assay and the ID_{50} of the D3 differentiation assay were entered into the statistical evaluation developed from the modified prediction model used by Scholz et al. 1999a:

$$\text{Log\_mean} = \frac{\lg IC_{50}D3 + \lg IC_{50}3T3 + \lg ID_{50}}{3}$$

$$D12\_3 = \frac{\lg IC_{50} \; \lg IC_{50}3T3}{2 - \lg ID_{50}}$$

Strong teratogenic/embryotoxic potential (Class III) – log_mean<0.5
Non-teratogenic/embryotoxic potential (Class I) – D12_3<0.25
Weak teratogenic/embryotoxic potential (Class II) – all remaining values

Contingency tables
Contingency tables were developed by Genschow et al. (2002) to allow the results of the EST to be evaluated and compared to *in vivo* results. Predictivity for

weakly (or strongly) embryotoxic substances is defined as an estimate of the likelihood that a positive prediction in the test correctly identifies a weak (or strong) embryotoxic substance under the proposed condition of use.

Precision for weak (or strong) embryotoxic substances is defined as the proportion of correctly classified weak (or strong) embryotoxic substances from the *in vitro* test versus chemicals that are weakly (or strongly) embryotoxic *in vivo*.

To evaluate the performance of the test the following criteria published by Genschow et al., 2002 can be used:

Chance = 33%
Insufficient < 65%
Sufficient ≥ 65%
Good ≥ 75%
Excellent ≥ 85%

3 Results and discussion

The EST with the prediction model developed by ECVAM has proven to have potential as a screening tool for chemicals. Pharmaceutical substances tend to be more biologically active than chemicals and therefore the suitability the EST to detect the embryotoxic potential of these substances was investigated.

All 6 chemicals tested previously and 7 of the 10 pharmaceutical substances produced by F. Hoffmann-La Roche AG were classified correctly. In addition, the

positive control 5-fluorouracil was also correctly classified as being strongly embryotoxic (the result was considered to be correct if the *in vitro* classification correlated with the *in vivo* classification). The endpoints and the resulting classifications of all the test substances can be found in Table 1. A contingency table and an evaluation of the classifications *in vitro* and *in vivo* are shown in Tables 2 and 3. Representative figures illustrating concentration response curves of a non-embryotoxic compound (R9) are shown in Figure 1 and Figure 2.

All of the non-embryotoxic substances were correctly identified, thus giving 100% precision (Tab. 3).

In this work, the EST and the developed prediction model produced an accuracy of 81% (Tab. 3). When evaluated in terms of the performance of the classification system developed by Genschow et al. (2002), this result would be regarded as "good" and a strong basis for the screening of pharmaceutical substances early in their development.

One of the three weakly embryotoxic substances was falsely classified as being strongly embryotoxic, leading to a precision of 66.7% (Tab. 3). R10 is a calcium antagonist, which caused anomalies in the heart and blood vessels. Its effect *in vivo* on the cardiovascular system suggests that the differentiation of embryonic stem cells into beating cardiomyocytes could be heavily impaired. It was evident

from the endpoint values that the effect was indeed much greater in the differentiation assay than on the viability of the D3 and 3T3 cells in the cytotoxicity assay, and this effect was sufficient to place the substance in a stronger class of embryotoxicity.

Five out of seven strongly embryotoxic substances were correctly identified, giving a precision of 71% (Tab. 3). Of the substances that were falsely classified, R2 is an enantiomer of R8, which was correctly classified *in vitro*. Both are analgesics. R2 caused cleft palate and resorptions *in vivo*. It had a clear effect on the inhibition of differentiation, but little effect on cell viability, thus classifying it overall as being weakly instead of strongly embryotoxic. In agreement with Scholz et al. (1999a), some substances that have no effect on the development of the heart could be correctly identified, but other treatments which affect particular tissues, including the nervous system or the palate, may not be detected.

The second falsely classified substance, R3 (a collagenase inhibitor), showed little difference between the endpoints of differentiation and cytotoxicity. This substance was found to cause severe craniofacial, cardiovascular and urogenital malformations *in vivo*. However, the effect on the cells was not significant and both D3 and 3T3 cells were equally affected. It may be that the compound requires metabolic activations,
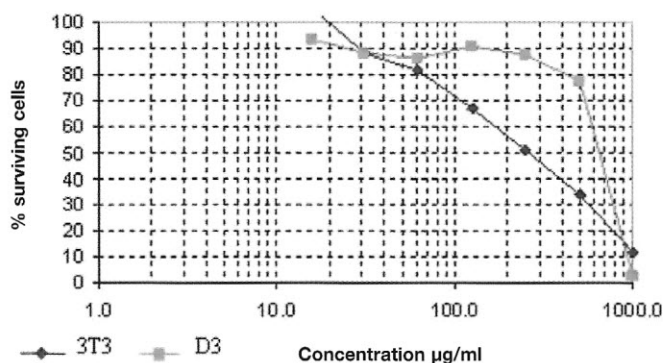

Fig. 1: Concentration response curves for the effect of R9 on the viability of D3 and 3T3 cells. The curves were obtained from two single experiments.
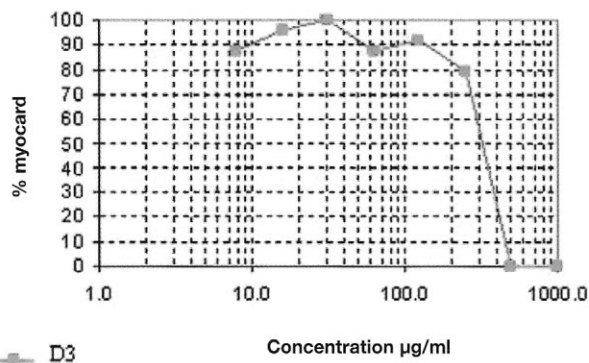

Fig. 2: Concentration response curves for the effect of R9 on differentiation of D3 cells. The curve was obtained from two single experiments.

which was not taken into consideration in these assays.

The EST has been developed using chemicals which require no metabolic processing to have effects. However, some *in vitro* experiments have been carried out with the inclusion of a metabolic activation system. Both Oglesby et al. (1986) and Piersma et al. (1991) added hepatocytes to the whole embryo culture system to test for embryotoxicity. It is known that thalidomide has little effect on cytotoxicity (Moreira et al., 1999) but in the presence of liver microsomes an anti-angiogenic effect is observed (Bauer et al., 1998; Sauer et al., 2000). Conversely, metabolic activation may reduce the toxic effect of substances (Piersma, 2004). Metabolic activation is yet to be explored extensively in the EST.

Typical patterns could often be found in the concentration-response curves. Strongly embryotoxic substances caused inhibition of growth and differentiation at low concentrations. Weakly and non-embryotoxic substances showed toxicity only at higher concentrations, whereby the differentiation potential tended to be the most strongly affected when substances that were weakly embryotoxic were tested; the cytotoxicity of 3T3 cells tended to be most affected when non-embryotoxic substances were tested (Scholz et al., 1999a; Genschow et al., 2002). A balanced combination of the endpoints of the cytotoxicity test with undifferentiated and differentiated cells and the differentiation assay is known to give the most accurate results (Scholz et al., 1999b).

In comparison to the ECVAM experiments, our results (accuracy of 81%) are a slight improvement on the prediction model used in the ECVAM study, which resulted in an accuracy of 78% (also classified as being "good") (Genschow et al., 2004). As already mentioned, all of those chemicals were known to require no any metabolic activation. Even though the pharmaceutical substances tested in this study were more active and R3 probably required metabolic activation, the accuracy obtained was slightly higher. Hence the model has potential for use in the pharmaceutical industry.

Tab. 1: Endpoints of the cytotoxicity and differentiation assays (mean ± SD) and their resulting classifications *in vitro* and *in vivo*.

$IC_{50}D3$ is the concentration of substance at which 50% of D3 cells have survived and $IC_{50}3T3$ is the measurement at which 50% of the 3T3 cells have survived. $ID_{50}D3$ is the concentration at which there is a 50% reduction in the differentiation of D3 cells into beating cardiomyocytes.

1 – non embryotoxic
2 – weakly embryotoxic
3 – strongly embryotoxic

	IC_{50} D3 (µg/ml) (n = 2)	IC_{50} 3T3 (µg/ml) (n=2)	ID_{50} (µg/ml) (n=2)	Classification	
				In vitro	*In vivo*
Penicillin G sodium salt	1000 ± 0	925 ±106	1000 ±0	1	1
Saccharin sodium salt	1000 ± 0	1000 ± 0	1000 ± 0	1	1
LiCl	550 ± 70	740 ± 57	140 ± 0	2	2
Boric acid	170 ± 57	729 ± 114	92 ± 42	2	2
Retinoic acid	0.0038 ± 0.0005	17.5 ± 0.7	0.00048 ± 0.0002	3	3
BrdU	0.71 ± 0.056	1.08 ± 0.45	0.48 ± 0.2	3	3
R6	9 ± 1	13 ± 2	18 ± 0	1	1
R7	26 ± 12	7 ± 2.8	24 ± 8.4	1	1
R9	700 ± 0	200 ± 70	300 ± 0	1	1
R8	50 ± 14	50 ± 14	29 ± 9.1	1	1
R3	42 ± 3.5	40 ± 0	40 ± 0	1	3
R2	67 ± 32	140 ± 7	1 ± 0.7	2	3
R1	4 ± 0	3 ± 1.4	2 ± 0.7	3	3
R10	2 ± 1	3 ± 1.2	1 ± 0.4	3	2
R4	2 ± 0.5	2 ± 0.8	0.0006 ± 0.0001	3	3
R5	2 ± 0.35	2 ± 0.35	0.032 ± 0.03	3	3

Tab. 2: Contingency table comparing *in vivo* and *in vitro* classifications

In vivo classification	No. of chems	No. of exp.	*In vitro* classification		
			1	2	3
1	6	6	6	0	0
2	3	3	0	2	1
3	7	7	1	1	5

Tab. 3: Summary of classifications for pharmaceutical substances and chemicals tested *in vivo* and *in vitro*.

Predictivity for non-embryotoxicity	85.7%
Predictivity for weak embryotoxicity	66.7%
Predictivity for strong embryotoxicity	83.3%
Precision for non-embryotoxicity	100%
Precision for weak embryotoxicity	66.7%
Precision for strong embryotoxicity	71.4%
Accuracy	81.3%

4 Conclusions and outlook

The EST can be carried out rapidly and is cheap. It is a true *in vitro* test using well-established cell lines, thus completely avoiding the use of animals. It was shown that the test is useful in the pharmaceutical industry as a screening tool very early in the development of substances to detect strongly embryotoxic substances and to remove them from further *in vivo* teratology testing. This would reduce the number of animals needed in embryotoxicity testing.

Further work has to be done to include a metabolic activation system and to increase the throughput of the assay. In addition, the potential use of the human embryonic stem cells that are now available should be explored.

References

Anon. (1999). INVITTOX Protocol 113. Embryonic Stem Cell Test (EST). Website http://ecvam-sis.jrc.it

Bauer, K. S., Dixon, S. C. and Figg, W. D. (1998). Inhibition of angiogenesis by thalidomide requires metabolic activation, which is species-dependent. *Biochem. Pharmacol. 55,* 1827-1834.

Doetschman, T. C., Eistetter, H., Katz, M. et al. (1985). The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J. Embryol. Exp. Morph. 87,* 27-45.

Genschow, E., Scholz, G., Brown, N. et al. (2000). Development of prediction models for three in vitro embryotoxicity tests in an ECVAM validation study. *In Vitro and Mol. Toxicol. 13,* 51-66.

Genschow, E., Spielmann, H., Scholz, G. et al. (2002). The ECVAM international validation study on in vitro embryotoxicity tests: results of the definitive phase and evaluation of prediction models. *ATLA 30,* 151-176.

Genschow, E., Spielmann, H., Scholz, G. et al. (2004). Validation of the embryonic stem cell test in the international ECVAM validation study on three in vitro embryotoxicity tests. *ATLA 32,* 209-244.

Laschinski, G., Vogel, R., Spielmann, H. (1991). Cytotoxicity test using blastocyst-derived euploid embryonal stem cells: a new approach to in vitro teratogenesis screening. *Repro. Toxicol. 5,* 57-64.

Moreira, A. L., Friedlander, D. R., Shif, B. et al. (1999). Thalidomide and a thalidomide analogue inhibit endothelial cell proliferation. *J. Neurooncol. 43,* 109-114.

Newall, D. R. and Beedles, K. E. (1994). The stem cell test – a novel in vitro assay for teratogenic potential. *Toxic. In Vitro 8,* 697-701.

Oglesby, L. A., Ebron, M. T., Beyer, P. E. et al. (1986). Co-culture of rat embryos and hepatocytes: in vitro detection of proteratogen. *Teratogen. Carcinogen. Mutagen 6,* 129-138.

Piersma, A. H., van Aerts, L. A., Verhoef, A. et al. (1991). Biotransformation of cyclophosphamide in post-implantation rat embryo culture using maternal hepatocytes in co-culture. *Pharmacol. Toxicol. 69,* 47-51.

Piersma, A. H. (2004). Validation of alternative methods for developmental toxicity testing. *Toxicol. Letters 149,* 147-153.

Rohwedel, J., Guan, K., Hegert, C. and Wobus, A. M. (2001). Embryonic stem cells as in vitro models for mutagenicity, cytotoxicity studies: present status and future perspectives. *Toxicol. In Vitro 15(6),* 741-753.

Sauer, H., Günther, J., Hescheler, J. and Wartenberg, M. (2000). Thalidomide inhibits angiogenesis in embryoid bodies by the generation of hydroxyl radicals. *Am. J. Pathol. 156,* 151-158.

Scholz, G., Pohl, I., Genschow, E. et al. (1999a). Embryotoxicity screening using embryonic stem cells in vitro: correlation to in vivo teratogenicity. *Cells Tissues Organs 165,* 203-211.

Scholz, G., Genschow, E., Pohl, I. et al. (1999b). Prevalidation of the embryonic stem cell test (EST) – a new in vitro embryotoxicity test. *Toxicol. In Vitro 13,* 675-681.

Seiler, A., Visan, A., Buesen, R. et al. (2004). Improvement of an in vitro stem cell assay (EST) for developmental toxicity by establishing molecular endpoints of tissue-specific development. *Repro. Toxicology 18,* 231-240.

Seiler, A., Buesen, R., Visan, A., and Spielmann, H. (2006). Use of Murine Embryonic Stem Cells in Embryotoxicity Assays: The Embryonic Stem Cell Test. In Methods in Molecular Biology: Embryonic Stem Cells- II, Edited by: K. Tuksen, Humana Press, Totowa, NJ, USA. 329, 371-375.

Spielmann, H., Pohl, I., Döring, B. et al. (1997). The embryonic stem cell test, an in vitro embryotoxicity test using two permanent mouse cell lines: 3T3 fibroblasts and embryonic stem cells. *In vitro Toxicol. 10,* 119-127.

Williams R. L., Hilton, D. J., Pease, S. et al. (1988). Myeloid leukemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature 336,* 684-687.

Wobus, A. M., Wallukat, G. and Hescheler, J. (1991). Pluripotent mouse embryonic stem cells are able to differentiate into cardiomyocytes expressing chronotropic responses to adrenergic and cholinergic agents and Ca^{2+} channel blockers. *Differentiation 48,* 173-182.

Wobus, A. M. and Boheler, K. R. (2005). Embryonic stem cells: prospects for developmental biology and cell therapy. *Physiol Rev. 85(2),* 635-78.

Correspondence to
Dr. Nicole Clemann
F. Hoffmann-La Roche Ltd.
Pharma Research
Bldg. 73/115b
Grenzacherstrasse 124
4070 Basel
Switzerland
Tel: +41-(0)-61 687 1949
Fax: +41-(0)-61 688 8101
e-mail: nicole.clemann@roche.com