Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Pharmaceutical and Biomedical Analysis 36 (2004) 285–294
JOURNAL OF PHARMACEUTICAL AND BIOMEDICAL ANALYSIS
www.elsevier.com/locate/jpba

Validation of the COSTIM bioassay for dendritic cell potency

Gopi Shankar*, Marc S. Fourrier, Michael A. Grevenkamp, Patricia A. Lodge

Northwest Biotherapeutics, Inc., 22322, 20th Avenue S.E., Bothell, WA 98021, USA

Received 7 March 2004; received in revised form 26 May 2004; accepted 28 May 2004
Available online 31 July 2004

Abstract

Dendritic cells (DCs) are increasingly prepared in vitro for use in clinical trials of human disease. Their utility in experimental immunotherapy has driven significant advances in the manufacture of these cells. Thus it has become imperative that, in concert with other quality control measures, a potency test be utilized for the GMP/GLP lot-release of DC products for preclinical and clinical studies. For this purpose we developed a novel method named the 'COSTIM bioassay', which selectively measures co-stimulatory activity, or functional potency of the DCs. In this method, T-cells stimulated with a sub-optimal amount of anti-CD3 antibody are unable to proliferate unless a source of co-stimulation (DCs) is added to the culture. We describe our validation of this method in this paper.

Keywords: Dendritic cells; Co-stimulation; Bioassay; Potency; Validation

1. Introduction

Dendritic cells (DCs) are increasingly prepared in vitro for use in the experimental immunotherapy of human disease, particularly cancer [1]. The production of DCs with optimal immunostimulatory properties and their exploitation in immunotherapy necessitated an understanding of the biology of these cells during in vitro culture. Therefore, various protocols for the culture of these cells were established with the specific purpose of human clinical use, and various advantages ascribed to each [2–8]. Recent improvements on DC manufacture include the use of serum-free media [9,10], the use of IL-13 (instead of IL-4) in concert with GM-CSF [11], or other cytokine combinations [12], the development of fully closed purification and culture systems [6,7,13], and novel DC maturation methods suitable for clinical application [14–16]. In light of such developments and the probability that different methods will produce distinct DCs with varying characteristics or heterogeneous populations, a critical need has emerged for the quality control of ex vivo manufactured DCs.

DC-based vaccines are unique products since DC lots (or batches) are patient-specific, unlike currently marketed biotherapeutic and pharmaceutic drugs. Leukocytes (generally, peripheral blood derived mononuclear cells) from a patient are obtained, cultured ex vivo, often exposed to a target antigen, and then infused back into the same patient with the expectation that a target-specific immune response will be stimulated in the patient's body resulting in a therapeutic effect. For preclinical toxicology and clinical studies, manufacturers of DCs not only need to assess the morphological and functional quality of each batch of these cells, but also be able to monitor the stability of these cells upon long term cryostorage. Quality control is generally achieved by phenotypic and functional testing of DCs. The mixed lymphocyte reaction (MLR) assay has served as the 'gold standard' for evaluating the functional ability of antigen presenting cells [17,18]. However, MLR assays take several days to complete and are non-representative of the activity of dendritic cells in context of their stimulatory interactions with autologous antigen-specific T-cells. Major histocompatibility (MHC) antigens constitute a powerful stimulus in the MLR, making it difficult to independently assess the impact of the co-stimulatory capacity of the DCs, that is, their antigen-independent potency. Furthermore, from the quality control perspective, the MLR is also a rather uncontrolled system since the stimulating alloantigens will vary

* Corresponding author. Present address: Centocor, Inc., 200 Great Valley Pkwy, Malvern, PA 19355, USA. Tel.: +1-610-407-8922.
E-mail address: gshanka3@cntus.jnj.com (G. Shankar).

0731-7085/$ – see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2004.05.025

vastly between the batches of T-cells (responder cells) and batches of DCs (stimulator cells). For example, when low MLR data are obtained for a sample, it is unclear whether they are the result of truly impotent DCs, or due to a significant histocompatibility match between the DCs and T-cells. Immunophenotyping by flow cytometric methods, which is been used to determine purity of DC preparations, has also been used as a surrogate of potency [19], but is obviously inaccurate because it does not measure cellular function; a highly pure but non-viable or apoptotic (and thus obviously impotent) DC preparation would erroneously pass such a potency test.

Given the limitations of MLR and immunophenotyping in determining DC potency, we developed a novel DC potency method named the 'COSTIM bioassay', which we demonstrate in this paper to be valid for 'lot-release' quality control testing. The COSTIM bioassay was developed to specifically measure co-stimulatory activity but not antigen processing and presentation, and also as an efficient and economical method for patient-specific lot testing. It is more relevant than MLR to the function of DCs in a vaccine product. Furthermore, the COSTIM bioassay is more accurate, reliable, and efficient than currently used DC potency tests. In this bioassay, T-cells are stimulated with a sub-optimal amount of anti-CD3 antibody, such that they remain unable to proliferate unless a source of co-stimulation (accessory cells, such as DC) is added to the culture. Thus, this bioassay is a functional test that selectively measures co-stimulatory activity, or functional potency, of DCs. The detailed protocol has been published elsewhere [20].

We developed the COSTIM bioassay for assuring the potency of our DC-based experimental vaccine for prostate cancer, DCVax^{TM}-Prostate (BCG-matured autologous DC loaded with prostate-specific membrane antigen). Since the development of a convenient, robust, accurate, and precise method is important for the 'lot-release' of a drug or vaccine, it was imperative that our method be validated. Assay validation at the Phase-I or II clinical trial stage is akin to an extensive characterization of the potency assay with the purpose of establishing that its performance characteristics are suitable and reliable for its intended analytical use. According to an FDA guidance document [21], this involves establishment of method validity through categorical characterization of the effects of a range of critical variables and assay conditions on assay performance (such as specificity, selectivity and robustness of the assay), and testing the established assay for its accuracy and precision against pre-determined acceptance criteria. By describing our validation of the COSTIM assay in this paper we intend to demonstrate that this assay is suitable for routine application in the quality control laboratory as a potency-determining test for dendritic cells. Based on our full validation, anyone intending to employ the COSTIM bioassay for other products may only need to perform a partial validation.

2. Materials and methods

All the work in this paper was performed under conditions of Current Good Manufacturing Practices (cGMP) in our quality control department and the validation parameters were in accordance with the Guidelines of the International Conference on Harmonization (ICH) [22,23].

2.1. Reagents and supplies

Bacillus Calmette-Guerin, commonly known as BCG (BCG-Tice®, Organon Teknika, Durham, NC), was inactivated in-house by heat and formalin-fixation prior to use for DC maturation. Interferon-γ (IFN-γ) (Actimmune®, InterMune Pharmaceuticals, Brisbane, CA) was used in the maturation of DC in some experiments, as described previously [16]. Monoclonal anti-human CD3 antibody, anti-CD54, anti-CD80, anti-CD86 (BD Pharmingen, San Diego, CA) and ^{3}H-thymidine (Perkin Elmer Life Sciences, Boston, MA) were other critical reagents used in this bioassay. Antibodies used were sterile, azide-free, and endotoxin-free analyte specific reagents (ASR), per the manufacturer's certificate of analysis.

2.2. Preparation of T-cell-enriched responder lymphocytes

Human peripheral blood mononuclear cells (PBMC) were isolated from freshly (within 24 h) leukapheresed blood via density gradient centrifugation using Ficoll (Sigma Chemical Co, St. Louis, MO). The PBMC were cryopreserved in 90% autologous serum and 10% dimethyl sulfoxide (DMSO) until further use. These PBMC were used either for DC culture or for preparing T-cells used in the bioassays. Enriched T-cells were prepared from allogeneic PBMC by negative depletion using anti-HLA-DR monoclonal antibody-conjugated paramagnetic beads (Dynal, Lake Success, NY). Thus, the PBMC were depleted of cells possessing potential co-stimulatory function. Each batch of enriched T-cells was tested for the presence of any remaining B-cells and monocytes, and these cells were found to comprise less that 0.5% after enrichment. This cell suspension consisted of 80–90% T-cells (the remainder being NK cells, which were previously found to be inert in this assay).

2.3. Preparation of purified monocytes, B-cells, and T-cells

Purification of monocytes, B-cells, and T-cells for use as stimulators was achieved by biomagnetic separation, using cell-specific paramagnetic bead preparations purchased from Dynal, and per the manufacturer's recommended protocol. The purity of each cell type after this enrichment was greater than 90%, as determined by flow cytometry.

2.4. Dendritic cells

Previously frozen PBMC were thawed in warm AIM-V medium (Life Technologies, Gaithersburg, MD), washed

once with PBS, and resuspended at (5–10) $\times$ 10^6 cells/ml in Opti-MEM™ medium (Invitrogen Corp., Carlsbad, CA) supplemented with 1% heat-inactivated autologous plasma. Ten millilitres of this cell suspension was dispensed into T-75 culture flasks and incubated for 1 h. After the incubation, non-adherent cells were resuspended and aspirated out, followed by stringent washing with cold PBS to remove loosely adherent cells. Fifteen millilitres of Opti-MEM™ medium containing 5% heat-inactivated autologous plasma, 500 IU/mL of rhGM-CSF (Leukine™, Immunex, Seattle, WA) and 500 IU/mL of rhIL-4 (R&D Systems, Minneapolis, MN) was added to each flask and the adherent cells (monocytes) cultured for 6 days. This method typically resulted in >90% $CD11c^+$ $HLA\text{-}DR^+$ "immature" DC that were >80% viable (determined by trypan blue exclusion), approximately 7% B-cells, and the remainder T-cells and NK cells. For maturation, these DC were treated with BCG alone or BCG plus IFN-γ for 24 h as described in the individual experiments.

2.5. The COSTIM bioassay

The detailed protocol is described in a separate publication [20]. Having found previously that both autologous and allogeneic T-cells proliferate equally in this assay, we used allogeneic T-cells in this assay. Previously cryopreserved DCs (stimulators) and T-cells (responders) were used in all experiments. Cells were thawed in warm AIM-V culture media, washed, and then resuspended at 1×10^5 live cells/mL (DCs) or 1×10^6 live cells/mL (T-cells). Cellular viability was above 80%. Dendritic cells (1×10^4 cells, in a 100 μL volume) were added to each of triplicate wells of a U-bottom 96-well plate (VWR International, West Chester, PA) followed by allogeneic T-cells (1×10^5 cells, in a 100 μL volume), and with or without 0.005 μg/mL anti-CD3 monoclonal antibody. The plate was incubated for 44 h in a humidified, 5% CO_2, 37 °C incubator. Tritiated (^{3}H)-thymidine (0.5 μCi per well, in 50 μL of AIM-V) was added to the culture wells for the last 18 h of the culture. Subsequently, cells were harvested onto glass fiber filters using a FilterMate™ harvester (Packard, Meriden, CT) and the incorporated radioactivity quantified in a TopCount™ scintillation counter (Packard). The result was calculated by subtracting the background MLR reactivity (T-cells and DC co-cultured in the absence of anti-CD3 antibody). For co-stimulatory molecule blocking studies sterile, azide-free, and low-endotoxin IgG_1 monoclonal antibodies specific for CD54, CD80, CD86, and an isotype control (BD Pharmingen) were added to the DCs at 1 μg/well for 1 h, prior to the addition of T-cells and anti-CD3.

2.6. Mixed lymphocyte reaction (MLR)

As in the COSTIM assay, previously frozen DCs (stimulators) and T-cells (responders) were used in all experiments. Cells were thawed in warm AIM-V media, washed, and resuspended at 1×10^5 (DCs) or 1×10^6 live cells/mL (T-cells). Cell viability was above 80%. Dendritic cells (1×10^4 cells, in a 100 μL volume) were added to each of triplicate wells of a U-bottom 96-well plate (VWR International, West Chester, PA) followed by T-cells (1×10^5 cells, in a 100 μL volume). The plate was incubated for 6 days in a humidified, 5% CO_2, 37 °C incubator. Tritiated (^{3}H)-thymidine (0.5 μCi per well, in 50 μL of AIM-V) was added to the culture wells for the last 18 h of the culture. After incubation, the cells were harvested onto glass fiber filters using a FilterMate™ harvester (Packard, Meriden, CT) and the incorporated radioactivity quantified in a TopCount(tm) scintillation counter (Packard).

3. Results and discussion

3.1. Specificity

Specificity is the property of a method to detect or quantify the target analyte. Monocytes and B-cells are capable of antigen presentation and express several cell-surface co-stimulatory molecules, albeit at lower levels than DCs. To demonstrate that the COSTIM assay specifically demarcates the higher magnitude of co-stimulatory capacity of these cell types, we enriched monocytes and B-cells (to greater than 90% purity) from one lot of PBMC by biomagnetic separation and used them as stimulators in the assay. In vitro cultured DCs, made from the same lot of PBMC, were purified to greater than 99% by fluorescent-activated cell sorting of CD14–$CD11c^+$ cells. As shown in Fig. 1A, monocytes were able to stimulate T-cell proliferation albeit at a much lower level than DCs. In contrast, B-cells possessed little or no co-stimulatory ability. To confirm the comparative difference in co-stimulatory capacities between DCs and monocytes, varying numbers of these stimulator cells were assessed in the COSTIM assay and the results depicted in Fig. 1B. At 5×10^4 stimulator cells per well, monocytes were indeed capable of significant co-stimulation of T-cell proliferation, although less than half of the response to an equal number of DCs. At 1.3×10^4 stimulator cells per well, monocytes provided negligible co-stimulatory activity (approx. 1400 CPM) in contrast to DCs (approx. 31278 CPM). Based on this, subsequent experiments used 1×10^4 DCs per well, unless otherwise specifically stated. These experiments clearly proved the specific ability of the COSTIM assay to elucidate DC-mediated co-stimulation of T-cell proliferation.

The T-cell proliferative response to DC-mediated co-stimulation in this bioassay is dependent on the co-stimulatory molecules on the DC cell-surface. To confirm this, immature and mature (BCG and IFN-γ co-treated for 24 h) DCs were compared in the bioassay. Varying ratios of immature or matured DCs and T-cells (1×10^5 per well) were co-cultured, as shown in Fig. 2A. In vitro matured DCs upregulate their expression of co-stimulatory molecules

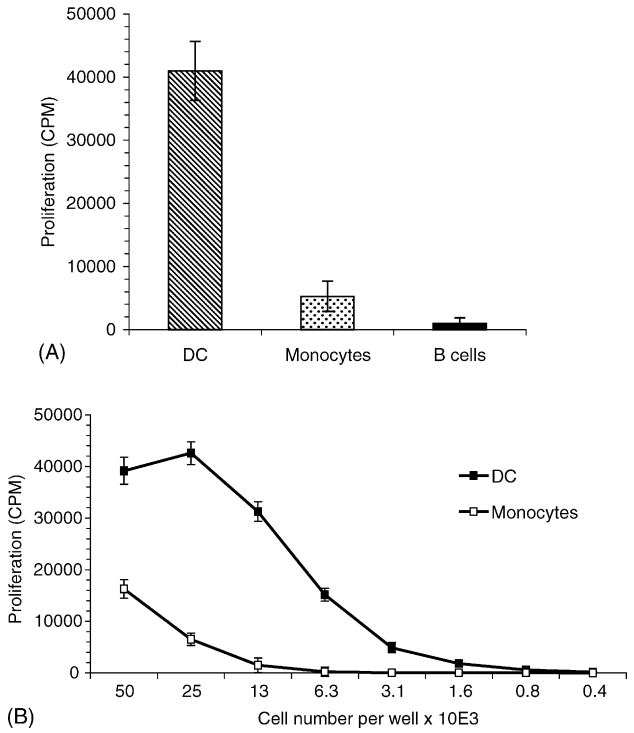

Fig. 1. Specificity of the COSTIM assay to accessory cells was demonstrated by (A) comparing equal numbers of DCs, monocytes, and B-cells, and (B) dose-response of varying numbers of DCs and monocytes per well.

[16]. Hence, as expected, there was a significantly higher response elicited by matured DCs than immature DCs. To address the issue of specificity to cell-surface co-stimulatory molecules at the molecular level, COSTIM cultures were set up and certain co-stimulatory molecules, or a combinations of them, were blocked using specific antibodies. Immature or matured DCs were exposed to anti-CD54, anti-CD80, anti-CD86, or an isotype-matched control blocking antibody for 1 h prior to the addition of T-cells. The results showed that when the cell-surface CD54, CD80, and CD86 were blocked during the culture, this response was significantly lowered (Fig. 2B). Blocking CD54, CD80, and CD86 together abrogated proliferation by over 95%, indicating that these three co-stimulatory molecules were responsible for practically all the co-stimulatory activity of DCs. This conclusion is strongly supported by the fact that the same three molecules were involved in the co-stimulation by both immature DCs and matured DCs. Similarly, blocking these molecules also inhibited over 75% of the IFN-γ secreted by the T-cells in the COSTIM culture (not shown), indicating that the same co-stimulatory molecules may be required for cytokine secretion by T-cells.

3.2. Selectivity

Selectivity is defined as the analytical specificity of an assay in the presence of potentially interfering substances that would normally be expected in the product. The ability of DC co-stimulated T-cell proliferation to remain unaffected by the presence of excipients was tested. T-lymphocytes, B-lymphocytes, and monocytes are expected excipients in ex vivo preparations of peripheral blood monocyte-derived DCs. To demonstrate selectivity, irradiated (20 Gy) allogeneic peripheral blood mononuclear cells (PBMC) were used as the source of excipients. The proportion of each excipient within PBMC was confirmed by flow cytometry (not shown) using antibodies against CD3 (T-cells), CD19 (B-cells), and CD14 (monocytes). 3×10^3 PBMCs (approx. 25% of the total cell number in the well) were added to 1×10^4 DCs in each well and co-cultured with T-cells in the COSTIM assay. The approximate proportion of 25% was chosen since we have a purity criterion of 75% DC for lot-release (in other words, a DC batch containing less than 75% $CD11c^+$ $HLA\text{-}DR^+$ cells fails). Fig. 3A shows the results of three independent experiments (runs) with three separate batches of DCs and donor-matched excipients. It was thus clear that addition of excess excipients did not interfere with detection in this assay.

Since monocytes were used for the preparation of DCs ex vivo, the effect of their presence in DC batches was investigated. Thus DCs (purified to greater than 99% by fluorescent-activated cell sorting of $CD14\text{–}CD11c^+$ cells from in vitro manufactured DCs), and monocytes (purified to greater than 99% by fluorescent-activated cell sorting of $CD14^+$ cells from PBMC) were mixed in different proportions (totaling 1×10^4 stimulator cells per well) and tested in the assay. As depicted in Fig. 3B, there was a predictable positive correlation between T-cell proliferation and the proportion of DCs, but a reciprocal negative correlation with the proportion of monocytes.

To get an estimate of systematic error in the COSTIM method, we took an alternative analytical approach. The data from the experiment depicted in Fig. 3B (mean of triplicate samples) are also shown in Table 1, where the recovery was calculated and shown to be in excess of 100%, with an average recovery of 112.8%. Thus there appeared to be an 12.8% level of systematic error in this bioassay.

3.3. Accuracy

Accuracy is a measure of the “trueness” of a method. In other words, it is the agreement between an experimentally measured value (from the assay) and an accepted reference (standard) or “true” value. Accuracy of a novel method can be estimated by comparing it with the results of another test method of known accuracy and precision (i.e., a “gold standard” method), or through reference material of known or generally accepted composition. No national or international reference materials exist for DCs.

Despite the fact that the COSTIM assay is functionally different from the MLR, we compared the two assays using one batch of DCs and T-cells. No pre-determined specifications were set for this experiment. In the MLR, DCs

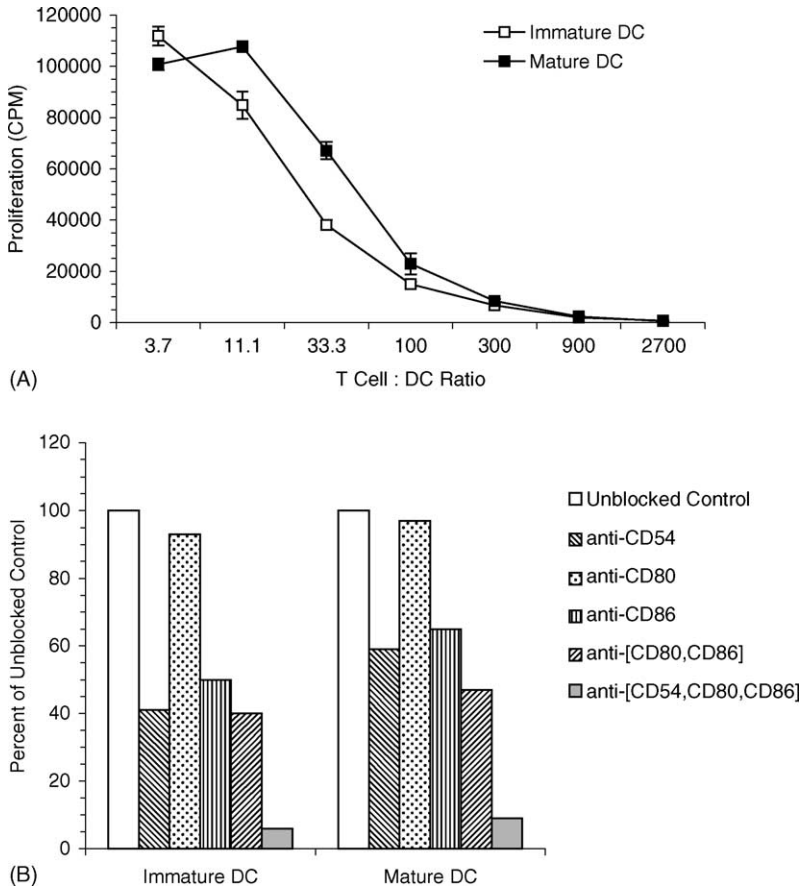

Fig. 2. Specificity of the COSTIM assay to cell-surface co-stimulatory molecules was demonstrated by (A) comparing the dose-response of varying numbers of immature and in vitro matured DCs, and (B) using blocking antibodies specific to the co-stimulatory molecules CD54, CD80, and CD86 to abrogate bioassay response.

and T-cells were co-cultured similar to the COSTIM assay, with the following exceptions: (a) no anti-CD3 was added to the cultures, and (b) the MLR incubation lasted 6 days. As shown in Fig. 4A, the response curves appeared parallel, except at the highest numbers of the DCs per well (indicated as the 3.7 and 11.1 T-cell:DC ratios in the figure). This was not surprising, since it is very likely that a higher T-cell:DC ratio in the culture required a shorter incubation period (i.e., less than 6 days) for capturing the optimal T-cell proliferation response. The high concentration of alloantigen present at such a high T-cell:DC ratio may have caused activation-induced cell death, resulting in the lower T-cell proliferation response. A 'comparison of methods plot' was generated for proliferation results between 1:33.3 and 1:2700 ratios of DCs to T-cells (Fig. 4B). The coefficient of determination (R^2) of 0.95 confirmed that the two methods provided similar results, albeit with a difference ('inter-method differential') of approximately 16,000 CPM

Table 1
Test of accuracy and selectivity

	Proportion of DC in culture (%)						
	100	80	60	50	40	20	0
Actual CPM	60588	53924	43649	36087	31611	16617	2569
Expected CPM[a]	–	51039	38922	32863	26804	14687	–
Percent response[b] (accuracy) (%)	100	85	68	55	48	23	–
Recovery[c] (selectivity) (%)	100	106	113	110	120	115	–

Varying proportions of DCs mixed with purified monocytes were tested in the COSTIM bioassay.

[a] The expected CPM was calculated as the multiple of 'proportion of DC in culture' and the 'actual CPM' obtained from the 100% pure DC group, plus the actual CPM obtained from the 0% DC (i.e., the 100% monocytes) group.

[b] Percent response was calculated for each proportion of DC as: 'actual CPM' of each group minus the 'actual CPM' obtained from the 0% DC (i.e., the 100% monocytes), as a fraction of the response of 100% pure DC group.

[c] Recovery of DC response was calculated by dividing 'percent response' by the 'proportion of DC in culture'.

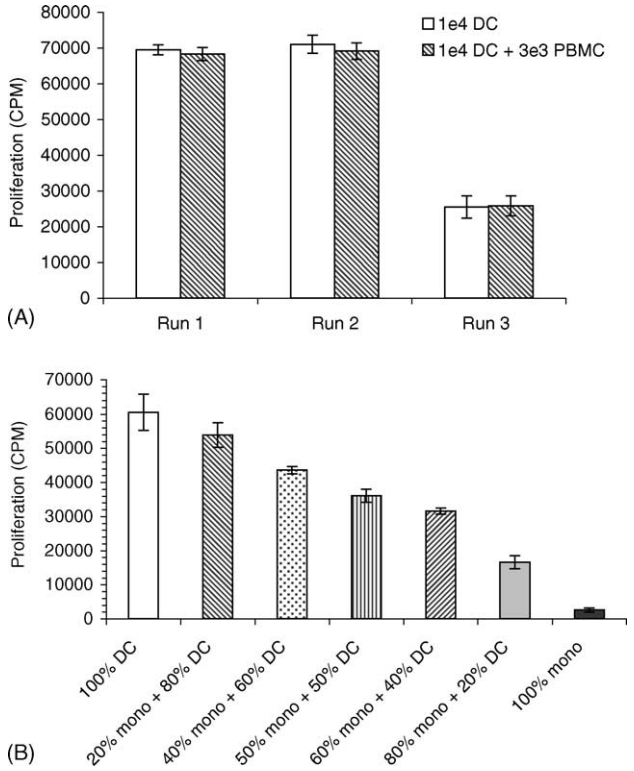

Fig. 3. Selectivity of the COSTIM assay to DCs in the presence of potential impurities was tested by (A) adding PBMC (to approximately 25% of the cell population) to pure DCs, and (B) adding various proportions of pure monocytes to pure DCs.

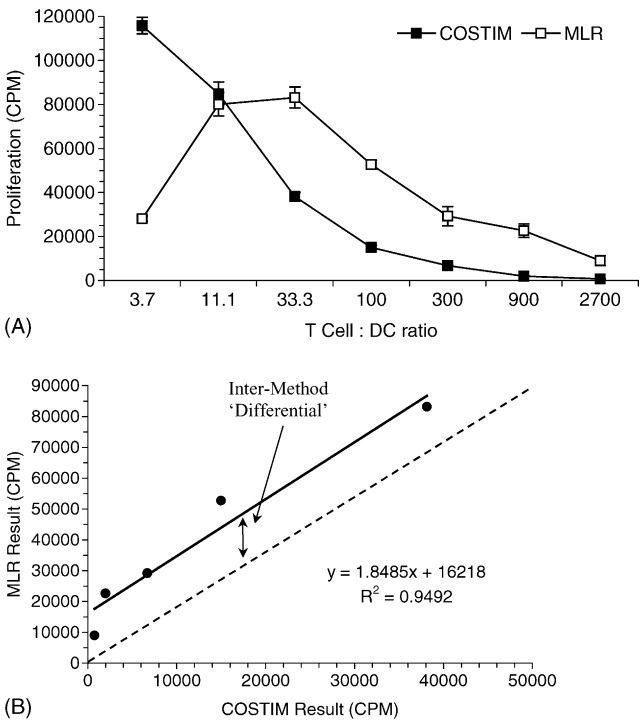

Fig. 4. Accuracy of the COSTIM assay was estimated by comparing it to the MLR (A). The two methods were compared in an alternate way to illustrate the inter-method differential in assay response (B).

between the two assays (see *y*-intercept). The dotted line represents the curve if both assays hypothetically produced identical results. A 'comparison of methods plot' should be considered carefully: the 16,000 CPM higher response seen in the MLR does not indicate that it is more sensitive than the COSTIM bioassay; the difference is merely due to the additional days of culture for the MLR. If the MLR is considered the true "gold standard" test of DC potency, then one might consider the distance between the two lines to represent the systematic error of this assay. We consider this difference an 'inter-method differential' since the MLR is not a true reference method nor, as discussed earlier, ideal for measuring DC potency. Nevertheless, this analytical approach indicated that the COSTIM assay could produce similar results as MLR, even though their mechanisms of actions vary significantly.

Alternatively, the data from the experiment depicted in Fig. 3B were analyzed to obtain an estimate of accuracy (Table 1). DCs were added in the indicated proportions to the culture, and the percent response (see table legend) was calculated from the proliferation CPM data obtained in that experiment. As seen in the table, the percent response was nearly identical to the proportion of DCs in the culture (the coefficient of determination, $R^2 = 0.99$). Thus, based on these experiments, we consider the COSTIM bioassay to be accurate for the determination of DC co-stimulatory potency.

3.4. Stability-indicating properties

We have previously found that cryopreservation of DCs did not cause loss of cell-surface co-stimulatory molecules (not shown). Furthermore, freshly prepared DCs and thawed (previously cryopreserved) DCs performed equally in the COSTIM bioassay and MLR (not shown). One aspect of DC stability that was of concern to us was their viability after variable periods of storage in liquid nitrogen, prior to preclinical or clinical use. To simulate cell death as a result of cryopreservation, DC suspensions were heat-killed by incubating them for 15 min in a 56 °C waterbath. This treatment killed all the DCs, as determined by PI staining and FACS analysis. Then calculated volumes of 99% viable DC and dead DC suspensions were mixed to obtain cell suspensions with pre-determined viability. These DCs were then placed in the COSTIM bioassay (1×10^4 total DC per well). As shown in Fig. 5, the proliferation response of T-cells in this bioassay decreased in direct proportion to the viability of DCs in the culture (coefficient of determination, $R^2 = 0.99$). Identical results were observed in a similar experiment using DCs killed by fixation with 1% paraformaldehyde for 15 min (not shown). Thus the COSTIM bioassay possesses stability-indicating properties.

3.5. Linearity

Without extensive clinical or preclinical data correlating DC potency with in vivo immunological activity, this

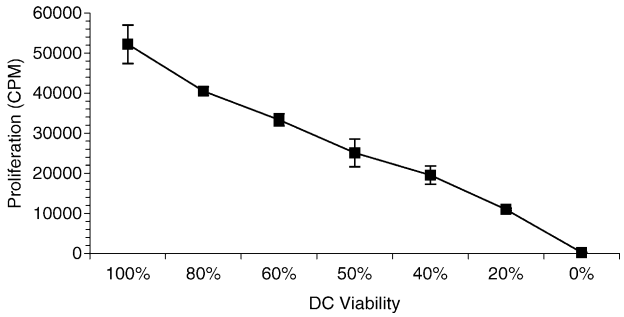

Fig. 5. Stability indicating property of the COSTIM assay was demonstrated by testing various levels of DC viability.

assay is considered quasi-quantitative. It is unknown what level of potency is physiologically needed, or relevant. Until clinical or preclinical data become available, we intend to use this assay only for determining whether DCs possess co-stimulatory activity. Thus linearity and a reportable range for this bioassay could not be defined at this point. However, elements of linearity were demonstrated in Figs. 1B, 2A and 3B, which indicated that the level of response in this assay is directly proportional to the total amount of co-stimulatory signal in the culture, be it the number of DCs per well, or maturation, that leads to increased total signal.

3.6. Precision

Precision is a measure of the degree of reproducibility of the analytical method under normal operating circumstances. In other words, it is the degree of agreement among individual test results when a procedure is applied repeatedly to multiple samplings of a homogenous batch of the product. Precision is measured mathematically by the random error, or imprecision, between replicate experiments. One batch each of DCs and T-cells was used to perform nine identical experiments: three analysts performed an experiment each on 3 consecutive days, and the resulting data were used to compute the different types of random experimental variation that comprise assay imprecision (Table 2).

Repeatability, or intra-assay precision, is the variation between replicate samples on the same plate (well-to-well variation within plate). We pre-determined the acceptance criterion for repeatability to be equal to 10%. The standard deviation (S.D.) and coefficient of variation (CV) between the triplicate samples on each day per analyst (shown in Table 2) were calculated and listed in Table 3. As shown, the intra-assay variation was very limited, with the average intra-assay CV under 5%. Thus, repeatability precision of this assay passed the acceptance criterion.

Table 3
Intra-assay precision, derived from the data shown in Table 2

	Day 1		Day 2		Day 3	
	S.D. (CPM)	CV (%)	S.D. (CPM)	CV (%)	S.D. (CPM)	CV (%)
Analyst 1	2533	4.0	1287	1.4	5852	8.4
Analyst 2	2336	3.6	1889	2.4	3478	6.1
Analyst 3	3320	5.1	2250	3.9	5608	7.3

Average intra-assay S.D. = 3172 CPM
Average intra-assay CV = 4.7%

Intermediate precision, comprising inter-assay precision (variation between runs on separate plates), inter-day precision (day-to-day variation within analyst), and inter-analyst precision (analyst-to-analyst variation), was also derived from the raw data in Table 2. We pre-determined the acceptance criterion for each type of intermediate precision to be equal to 20%. To determine inter-assay precision, means of the triplicate samples were first calculated (Table 4). The S.D. and CV between the mean results on each day per analyst were then calculated. As expected, the inter-assay imprecision (average CV of 16.7% and S.D. of 11,593 CPM) was higher than intra-assay precision. Inter-day precision was determined by calculating the S.D. and CV between the 3 days for each analyst. The results showed that inter-day precision was indifferent from inter-assay precision, since the average inter-day CV was 17.2%, and the S.D. was 12,049 CPM. Finally, inter-analyst precision was also determined by calculating the S.D. and CV between the three analysts on each day. The results showed that inter-analyst precision was also similar to the other two types of intermediate precision, with the average inter-analyst CV being 13.2%, and the S.D. being 9590 CPM. Thus, intermediate precision of this assay passed the acceptance criteria.

The precision of the COSTIM assay is, therefore, well within expectation for a cell-based bioassay.

3.7. Robustness

Robustness is an indication of the reliability of an assay, assessed by the capacity of the assay to remain unaffected by small, but deliberate, variations in method parameters.

Table 2
Test of precision

	Individual sample results (CPM)								
	Day 1			Day 2			Day 3		
Analyst 1	61938	65966	61292	94838	92819	92445	72515	73237	62759
Analyst 2	64329	67379	62789	76825	77524	80389	57327	60677	53722
Analyst 3	66213	67251	61052	59015	58254	54794	80186	78866	69881

Table 4
Intermediate precision, derived from the data shown in Table 2

	Mean result of triplicate samples (CPM), $n = 9$			Inter-day precision, $n = 3$		
	Day 1	Day 2	Day 3	S.D. (CPM)	CV (%)	
Analyst 1	63065	93367	69504	15964	21.2	Average inter-day S.D. = 12049 CPM
Analyst 2	64832	78246	57242	10636	15.9	Average inter-day CV = 17.2%
Analyst 3	64839	57354	76311	9548	14.4	
Average inter-assay S.D. = 11593 CPM						
Average inter-assay CV = 16.7%						
	Inter-analyst precision, $n = 3$					
S.D. (CPM)	1022	18083	9664			
CV (%)	1.6	23.7	14.3			
Average inter-analyst S.D. = 9590 CPM						
Average inter-analyst CV = 13.2%						

Three parameters that we believed were critical to robust assay performance were the concentration of anti-human CD3 antibody in the culture, the incubation period, and changes in T-cell lots.

Since the COSTIM assay protocol specified the use of 0.005 μg/mL of the antibody we tested assay performance at the 80 and 120% levels of this concentration (0.004 and 0.006 μg/mL). Based on our pipetting steps, the dilutions performed with the antibody, and the imprecision of the pipettors used (from the manufacturer's certificate of analysis of the instruments), these variations in final antibody concentration were deemed possible during actual runs. Three separate lots of DCs were tested using one responder T-cell lot. As shown in Table 5, the variation due to changes in antibody concentration for each DC lot was negligible (average CV for the three DC lots was 3.9%), well within the precision of this assay. The protocol also specifies a 44 h total incubation of the cell culture, including the 18 h incubation with tritiated-thymidine. We tested the effect of varying the incubation times by 1 h over or under. As shown in Table 6, we tested the combinations of the post-tritiated-thymidine period and the total culture incubation period using three separate lots of DCs and one T-cell lot. The results showed that this variation was also low (average CV for the three DC lots was 12.2%), well within the precision of this assay. Finally, we tested six different lots of responder T-cells against three DC lots using the specified COSTIM assay protocol. As indicated in Table 7, the variation was rather high (average CV for the three DC lots was 30.6%), but not at an unacceptable level for a cell-based bioassay. With prudent up-front qualification of T-cell lots, this variation could potentially be lowered.

Table 5
Robustness—effects of variation of anti-CD3 antibody concentration

Antibody conc. (μg/ml)	DC lot #1	DC lot #2	DC lot #3
0.004	56304	82469	28667
0.005	61683	77388	26761
0.006	57972	79495	26929
S.D. (CPM)	2753	2553	1055
CV (%)	4.7	3.2	3.8

3.8. Estimation of a lot-release 'cut-point' specification for DC potency

The minimal potency criterion for DCs was initially estimated by a calculation using the results of co-stimulatory

Table 6
Robustness—effect of variations in incubation periods

Total incubation time (h)	^{3}H-Thymidine incubation time (h)	DC lot #1	DC lot #2	DC lot #3
43	17	54645	36375	31259
44	17	56499	37185	35477
45	17	49685	52653	43737
43	18	58932	42682	28867
44	18	58968	46188	40244
45	18	66991	43615	39548
43	19	53498	44502	31137
44	19	54814	43831	39249
45	19	66407	42132	44765
S.D. (CPM)		5768	4805	5738
CV (%)		10.0	11.1	15.4

Table 7
Robustness—effect of changing T-cell lots

T-cell lot	Mean result of triplicate samples (CPM)		
	DC lot #1	DC lot #2	DC lot #3
T-cell lot #1	72281	64784	72995
T-cell lot #2	60545	60647	66040
T-cell lot #3	68757	64608	64479
T-cell lot #4	52658	42645	50244
T-cell lot #5	91772	75369	91192
T-cell lot #6	34116	26568	37412
S.D. (CPM)	19461	17845	18549
CV (%)	30.7	32.0	29.1
Average inter-lot S.D. = 18618 CPM			
Average inter-lot CV = 30.6%			

potency of monocytes. This was determined from the results of three different batches (donors) of pure monocytes, run in triplicate wells of the culture plate. The average potency CPM result and S.D. for 100% monocytes were 5287 and 2384, respectively (Fig. 3B). With a confidence of greater than 99%, monocytes were expected to stimulate a proliferation of less than or equal to the mean plus 3 S.D. units, which equaled 12,439 CPM. Since T-cells alone were allowed to proliferate up to 2000 CPM in this assay, adding it to the above-determined number resulted in 14,439 CPM. Thus, 15,000 CPM was initially considered our pass/fail 'cut-point' specification for DCs and was later supported by an alternative approach, as described below.

We employed the COSTIM method to test the potency of our prostate-specific membrane antigen (PSMA)-loaded dendritic cell product, DCVax^{TM}-Prostate, which was used in a phase-I/II trial for the treatment of prostate cancer patients. Twenty-three lots of DCVax^{TM}-Prostate were tested, of which most were used in the trial and some were preclinical pilot or control lots. The purpose of including these data in the validation were two-fold: (1) to confirm that under the real, manufacturing and operating environment, the COSTIM assay would produce results similar to prior observations, and (2) to re-estimate the lot-release 'pass/fail' specification cut-point for DCs in this assay.

As shown in Table 8, the MLR response for all lots tested was low (range 302–4746 CPM), expectedly since this assay largely precluded alloantigenic MLR responses due to the short culture period. The range of COSTIM potencies was 20,847–73,144 CPM, the mean was 39,914 CPM and the S.D. was 15,896 CPM. The potency data were normalized by logarithmic conversion, and then the mean and S.D. were calculated in order to derive an estimated pass/fail specification 'cut-point' for DCs in this assay. To be conservative it is generally advised to allow a feasible maximum level of false-negatives in a potency test, so as to minimize the false-positives. Allowing a 5% false-negative rate (mean − 1.645 S.D.) the cut-point would be at 18,789 CPM, and allowing a 2.5% false-negative rate (mean − 1.96 S.D.) the cut-point would be at 16,489 CPM. Allowing a 0.5% false-negative rate (mean − 2.576 S.D.) the cut-point would be at 12,774 CPM, which would be too risky as potent monocytes could erroneously pass this test. Our initial specification of 15,000 CPM allowed for a 2.2% false-negative rate, which was acceptable and therefore retained as the lot-release 'cut-point' specification for the potency of our DC-based product.

Table 8
Implementation of the COSTIM test on samples

DCVax^{TM}-Prostate lot #	Background MLR (T-cells + DC)	COSTIM potency (proliferation − MLR)	$\log_{10}$ COSTIM potency
1	627	19224	4.2838
2	497	35490	4.5501
3	700	38942	4.5904
4	2813	20847	4.3190
5	812	41428	4.6173
6	355	23025	4.3622
7	1569	47941	4.6807
8	1468	65242	4.8145
9	1058	52413	4.7194
10	3813	56685	4.7535
11	1432	73144	4.8642
12	3586	26049	4.4158
13	3324	45908	4.6619
14	665	32254	4.5086
15	274	27632	4.4414
16	302	22656	4.3552
17	774	52954	4.7239
18	484	27108	4.4331
19	632	28038	4.4477
20	1395	50952	4.7072
21	327	24328	4.3861
22	1242	65132	4.8138
23	3932	40622	4.6088
Mean	1395	39914	4.5678
Range	302–4746	20847–73144	
S.D. (CPM)	1210	15896	0.1752

4. Conclusions

We have demonstrated herein that the COSTIM assay is dependent on the co-stimulatory activity of accessory cells. It is a specific, selective, and accurate bioanalytical method that elucidates DC potency. We have demonstrated that the method is robust, and that the levels of imprecision in this assay are within reasonable limits for a cell-based bioassay. Based on this work, we consider the COSTIM assay valid for use as a lot-release potency test for our DC-based products.

Acknowledgements

The authors thank Mr. Daniel Wilds and Dr. Alton Boynton for sponsoring this work, and Mr. Chad Greenfield and Ms. Kristin Herndon for their technical assistance.

References

[1] F.O. Nestle, J. Banchereau, D. Hart, Nat. Med. 7 (2001) 761–765.
[2] G. Schuler, N. Romani, Adv. Exp. Med. Biol. 417 (1997) 7–13.
[3] B. Thurner, C. Roder, D. Dieckmann, M. Heuer, M. Kruse, A. Glaser, P. Keikavoussi, E. Kampgen, A. Bender, G. Schuler, J. Immunol. Meth. 223 (1999) 1–15.
[4] B. Feuerstein, T.G. Berger, C. Maczek, C. Roder, D. Schreiner, U. Hirsch, I. Haendle, W. Leisgang, A. Glaser, O. Kuss, T.L. Diepgen, G. Schuler, B. Schuler-Thurner, J. Immunol. Meth. 245 (2000) 15–29.
[5] A. Bender, M. Sapp, G. Schuler, R.M. Steinman, N. Bhardwaj, J. Immunol. Meth. 196 (1996) 121–135.
[6] V. Pullarkat, R. Lau, S.M. Lee, J.G. Bender, J.S. Weber, J. Immunol. Meth. 267 (2002) 173–183.

[7] S. Tuyaerts, S.M. Noppe, J. Corthals, K. Breckpot, C. Heirman, C. De Greef, I. Van Riet, K. Thielemans, J. Immunol. Meth. 264 (2002) 135–151.
[8] B. Obermaier, M. Dauer, J. Herten, K. Schad, S. Endres, A. Eigler, Biol. Proced. Online 5 (2003) 197–203.
[9] K. Tarte, G. Fiol, J.F. Rossi, B. Klein, Leukemia 14 (2000) 2182–2192.
[10] B. Goxe, N. Latour, M. Chokri, J.P. Abastado, M. Salcedo, Immunol. Invest. 29 (2000) 319–336.
[11] L. Piemonti, S. Bernasconi, W. Luini, Z. Trobonjaca, A. Minty, P. Allavena, A. Mantovani, Eur. Cytokine Netw. 6 (1995) 245–252.
[12] T. Buchler, R. Hajek, L. Bourkova, L. Kovarova, R. Musilova, A. Bulikova, M. Doubek, A. Svobodnik, I. Mareschova, P. Vanova, E. Tuzova, P. Vidlakova, J. Vorlicek, M. Penka, Vaccine 21 (2003) 877–882.
[13] R.V. Sorg, Z. Ozcan, T. Brefort, J. Fischer, R. Ackermann, M. Muller, P. Wernet, J. Immunother. 26 (2003) 374–383.
[14] A.W. Lee, T. Truong, K. Bickham, J.F. Fonteneau, M. Larsson, I. Da Silva, S. Somersan, E.K. Thomas, N. Bhardwaj, Vaccine 20 (2002) A8–A22.
[15] C. Boccaccio, S. Jacod, A. Kaiser, A. Boyer, J.P. Abastado, A. Nardin, J. Immunother. 25 (2002) 88–96.
[16] G. Shankar, L.A. Pestano, M.L. Bosch, J. Transl. Med. 1 (2003) 7.
[17] R.M. Steinman, M.D. Witmer, Proc. Natl. Acad. Sci. USA 75 (1978) 5132–5136.
[18] W.C. Van Voorhis, J. Valinsky, E. Hoffman, J. Luban, L.S. Hair, R.M. Steinman, J. Exp. Med. 158 (1983) 174–191.
[19] T. Monji, B. Smits-Sana, L. Qian, L. Stark, J. Therond, C. Shiomoto, M.V. Peshwa, Development of assay for assessment of dendritic cell product potency, Presented at the International Society for Hematotherapy and Graft Engineering (ISHAGE), San Diego, CA, 2000.
[20] G. Shankar, R. Bader, P.A. Lodge, J. Immunol. Meth. 285 (2004) 293–299.
[21] Guidance for Industry: Bioanalytical Method Validation, US Department of Health and Human Services, Food and Drug Administration (CDER and CVM), May 2001.
[22] Proceedings of the International Conference on Harmonization (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use, Section Q2A: Text on Validation of Analytical Procedures, March 1995.
[23] Proceedings of the International Conference on Harmonization (ICH) of Technical Requirements for Registration of Pharmaceuticals for Human Use, Section Q2B: Validation of Analytical Procedures: Methodology, May 1997.