Points for Discussion in the Selection of Dose Levels for Mammalian *in vivo* Tier 1 assays

The purpose of the mammalian *in vivo* Tier 1 assays is to provide information on the potential of agents to interfere with various aspects of the endocrine system and to use pharmacological responses to detect potential activity. These assays are primarily aimed at the detection of potential endocrine activity and as such EDSTAC proposed the view that the test systems employed should be subject to reasonably high dose levels, since the objective of the assays was not to generate detailed dose response information to be used in the risk assessment of the agents under test (which would be provided in Tier2 assays).

Maximally Tolerated Dose (MTD)

The EDSTAC also endorsed the concept that a maximally tolerated dose (MTD) should be employed with any other dose levels used at a fraction of the MTD. Also endorsed, was the concept of a limit dose level (1000 mg/kg/d orally) such that *in vivo* studies would not be conducted at ever increasing dose levels until toxicity was observed, rather that an upper bound could be set that would be acceptable to show that the test agent had been adequately evaluated.

Dose selection for any *in vivo* study should be based on all the information available on the agent to be tested and include previous toxicity data, predicted structure activity relationships, membership of a chemical class etc. The assignment of an MTD has long been a thorny issue for toxicologists. One approach would be to use precedent and employ the Agency's standard default position employed for studies of similar type or duration. For example, in reproductive and developmental studies this could be up to 10% mortality at the highest dose level tested. Even for screens of endocrine-like activity, this would be seen by most investigators as too severe. However, it is imperative that the various test systems utilized must be seen to be stressed, such that agents that have tested negative in Tier1 screens can confidently be placed in the "hold" box and require no further testing at this time and provide a potential registrant with some surety that the negative data obtained would be considered adequate by the Agency. The objective of the highest dose level employed would be to exert some overt systemic toxicity, but not lethality, to confirm the test system had undergone significant stress. The objective of the Tier1 *in vivo* assays is only to provide evidence of endocrine-like activity, not definitive dose response assessments (to be determined in Tier2) and thus a level of false positive response is acceptable. If the protocol used does not employ a limit dose, then a second dose level will normally be employed in the Tier assays. This second dose level should be biased towards the upper end of any dose response, since detection of activity is the prime purpose, but allow a clear separation of the seleccted dose level from the MTD. It is suggested that one half of the MTD be used for the setting of a second, lower dose level.

Toxicological indices likely to be of utility in assessing an MTD are those used most commonly by toxicologists and include changes in body weight, food consumption and clinical signs. As guidance it is suggested that a 10% decrease in terminal body weight, compared to control, be indicative of an MTD. Such a value would be less than the default EPA criteria, but high enough to account for the pharmacological responses of estrogens on appetite. However, the ability to

ascribe an MTD is not a "bright line." That is, it would be unreasonable to ask an investigator to repeat a study because the value obtained at the highest dose level was only a 9% decrease in body weight compared to the control and "therefore an MTD had not been achieved" and likewise, it would be unacceptable to dismiss any findings noted at an 11% decrease in body weight compared to the control, since this had "exceeded an MTD." Clearly MTD determinations would be based on individual chemicals and could be ascribed based on clinical signs or systemic target organ toxicity noted at necropsy if this was more appropriate than body weight. No substitute can be made for good judgement on the part of the toxicologist.

Use of Pilot Studies

Whilst the proposed mammalian *in vivo* assays are all to be conducted in the rat, they do span a range of other variables including gender, age, route of administration, duration of exposure and castration status. Thus durations of studies can vary from 3 days in the uterotrophic assay, 10 days in the Hershberger study, 20 days in the pubertal female 30 days in the pubertal male and greater than this for a potential *in utero*/ lactational study. In many instances dose levels may be selected based on available information in the rat by the route selected, for approximately the same period of exposure. However, in some circumstances little or no data will be available and a pilot study will be required in order to set dose levels for the assay. It is recommended that in the cases where minimal or no information is available that a stepwise approach be taken using the route of exposure to be employed in the assay and maximally to be for the duration of the study. Should adverse responses be noted, dosing should cease, potential recovery observed and then the dose level significantly reduced by at least a half log. Similarly, if no adverse responses are noted after several days, the dose level may be increased by a half log (or more) progressively. The maximum number of animals to be employed should be no more than four or five per group. In this manner a minimal number of animals will be employed and significant adverse responses and potential suffering of the animals be minimized.
A cautious approach to pilot studies will ensure reasonable information will be available on which to base the assay. In the case of the uterotrophic assay where the subcutaneous route is employed for only three days, it would also be sensible to also take and weigh the uterus and thus useful information could also be available from the pilot study.

In these studies the production of an increase in a pharmacologically important organ weight, even with body weight reductions (e.g. uterine weight in a pilot for a pubertal female assay or uterotrophic assay) may be sufficient to indicate that a positive response has occurred. However, significantly more care and inspection is required when decreases in organ weight are accompanied by large body weight reductions (e.g. prostate weight reduction in a pubertal male study)

Efforts in validation

In generating the guidance documents and protocols for dose setting and validation studies, it is recommended that these activities should be conducted in parallel. In order for a reasonable assessment of the variability and transferability of any protocol theis must be conducted with agents whee the dose levels are given. In this way the validation trials will be addressing two protocol issues, if dose levels are specified can individual laboratories produce comparable results. That is is the protocol robust and transferable and is the laboratory competent to undertake the assay specified by the protocol.

One test agent example employed for the interlaboratory validation study of any assay should be the investigation of an unknown agent in which the testing laboratory would be required to perform a dose setting exercise prior to undertaking the "definitive" assay. This activity should be undertaken in a parallel exercise. If the laboratories are given guidance on the conduct of a pilot study, competent laboratories should arrive at a selection of dose levels that can be scientifically justified (ie a test of the adequacy of the guidance and the competence of the scientists conducting the test in providing reasoned arguments based on data). If the scientific reasoning is sound on dose level selection, based on the information available to them, then the scientists conducting the assay in the trial will have adopted a correct approach and the dose levels will likely be qualitatively similar, if not quantitatively the same. Even so the ability to discern a response, the objective of the screens, should be fulfilled.