4. EVALUATION OF THE REPRODUCTIVE DEVELOPMENTAL TOXICITY DATA SET FOR DBP

This chapter presents the evaluation of the available toxicity data for the development of the male reproductive system following DBP exposure and the MOA(s) (see Chapter 2 and glossary for definition) that contribute to these outcomes. We used the compilation of the male reproductive toxicology literature cited in the 2006 external peer review draft IRIS Tox Review for DBP (U.S. EPA, 2006a) as a starting point for our toxicology literature review for this case study. Each toxicology study was examined for the lowest dose and low-incidence effects in order to identify the full spectrum of male reproductive developmental effects. In a second evaluation, we used available mechanistic information for each endpoint to identify potential MOAs. Endpoints with MOA information have support for phenotypic anchoring to some of the observed DBP gene expression changes (further discussed in Chapter 5). Endpoints with unexplained MOAs were used to identify and focus future research needs to study the mechanisms that underlie those endpoints using genomics and other techniques.

An extensive toxicological data set exists for DBP that includes acute and subchronic studies in multiple species, multigeneration reproduction studies in rodents, and studies that assess developmental outcomes following *in utero* or perinatal/postnatal exposures. Following DBP exposure during the critical stages of development, the male reproductive system development is perturbed in rodent studies (Gray et al., 1999, 2001; Mylchreest et al., 1998, 1999, 2000). Two MOAs of DBP, for a number of these outcomes, have been well established (David, 2006; Foster, 2005). The 2006 external draft IRIS Tox Review for DBP (U.S. EPA, 2006a) selected reduced fetal testicular T levels, observed in Lehmann et al. (2004), as the critical effect for the derivation of acute, short-term, subchronic, and chronic reference values for DBP. This case study evaluated information from genomic and other gene expression studies to target and further elucidate the molecular events underlying these developmental outcomes (see Chapter 5). The intent of performing this evaluation of the toxicology studies was to examine the usefulness of the toxicogenomic data in characterization of the MOA(s) that contribute to the adverse outcomes. We also examined the data for low-dose or low-incidence findings because such data may aid the interpretation of toxicological outcomes that can be misinterpreted as transient (e.g., AGD), or nonadverse due to low incidence or magnitude (e.g., not statistically

significant incidences of gross pathology findings in male offspring reproductive organs, or alteration of fetal T levels).

4.1. CRITERIA AND RATIONALE FOR INCLUSION OF TOXICOLOGY STUDIES IN THE EVALUATION

Figure 4-1 illustrates the process for evaluating the DBP toxicology data set for the case study (Section 4.2 discusses the later steps of the evaluation process in more detail). The first step in the process was the identification of studies to be included for consideration in the case study. We identified a number of study selection criteria in Step 1. One criterion of prime importance was that the studies should include exposures to DBP during sensitive periods of male reproductive system development. Secondly, a no-observed-effect level (NOEL), lowest-observed-effect level (LOEL), or BMDL would need to be identified for presumably adverse outcomes in the reproductive organs and/or function of male offspring. Additionally, the studies would need to be of adequate quality in order to establish confidence in the study conduct, methods, and results. These criteria, taken together, define a subset of the available toxicology studies that were considered possible candidates for determining the POD for derivation of reference values of various exposure durations in the 2006 external peer review draft Tox Review for DBP (see Tables 4-1, 4-2, and 4-3 in U.S. EPA, 2006a). These candidate study lists were considered during the external peer review of the IRIS document, conducted in July 2006, thereby providing a measure of confidence in their inclusiveness and veracity for the purpose of this case study. Though there are observable adverse effects on male reproductive system development in multiple species, the only available and relevant genomic studies with DBP (i.e., those that addressed effects on male reproductive system development following prenatal exposures) were conducted in rats. Table 4-1 lists the studies that were identified for inclusion as of July 2006. For each study, the following information was summarized: a description of the dose and exposure paradigm, the treatment-related outcomes observed at each dose level, and the experimentally derived reproductive NOEL and/or LOEL. The terms NOAEL and LOAEL are not used in this case-study report, although these terms are commonly used in risk assessment, because some study reports do not address the issue of adversity of observed study outcomes. In addition, some study reports do not specifically define NOELs or LOELs. For that reason, Table 4-1 presents those outcomes that could be considered biomarkers

of effects on the male reproductive system that were reported by the study authors, without specific consideration or judgment of adversity.

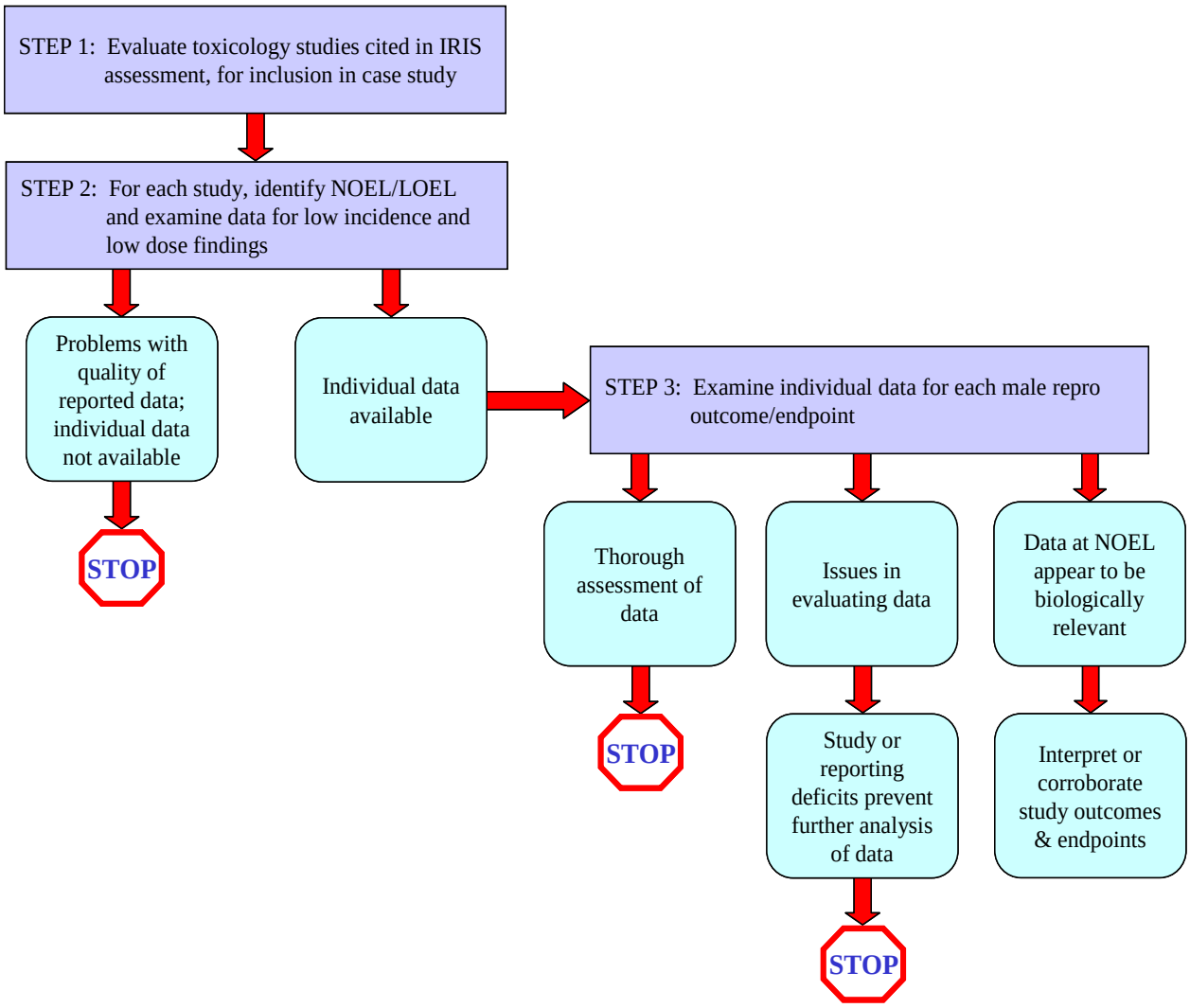

Figure 4-1. The process for evaluating the male reproductive developmental toxicity data set for low-dose and low-incidence findings. IRIS assessment, the 2006 external peer review draft IRIS Tox Review for DBP.

Table 4-1. (continued)

Study[a]	Species (strain), duration, and exposure	Reproductive system effects	NOEL mg/kg-d	LOEL mg/kg-d
Fisher et al., 2003	Rat (Wistar); GDs 13−21; 0 or 500 mg/kg-d	Cryptorchidism, hypospadias, infertility, and testis abnormalities similar to human testicular dysgenesis syndrome; abnormal Sertoli cell-gonocyte interaction.		500
Gray et al., 1999	Rat (Long-Evans) (P0); PND 21—adult; 0, 250, 500, or 1,000 mg/kg-d	At 250, 500, and 1,000 mg/kg-d, delayed puberty. At 500 and 1,000 mg/kg-d, reduced fertility related to testicular atrophy and reduced cauda epididymal sperm numbers.		250
	Rat (Long-Evans) (F1); GD 0−PND 21; 0, 250, 500, or 1,000 mg/kg-d	At 250 and 500 mg/kg-d, reproductive malformations (low incidences of hypospadias, testicular nondescent, and uterus unicornous); reduced fecundity.		250
	Rat (Long-Evans) (F1); GD 14 to PND 3; 0 or 500 mg/kg-d	Reduced AGD, retained nipples, permanently reduced androgen-dependent tissue weights.		500
Kim et al., 2004 **Ab**	Rat (SD); GDs 10−19; 0, 250, 500, or 700 mg/kg-d	Decreased testes and accessory sex organ weight; delayed testis descent; increased expression of estrogen receptor in testes.		250 (presumed) [b]
Kleymenova et al., 2004 **Ab**	Rat (strain not specified); GDs 12−17, 19, 20; 0 or 500 mg/kg-d	Altered proliferation of Sertoli and peritubular cells; multinucleated gonocytes; changes in Sertoli cell-gonocyte interactions.		500
Kleymenova et al., 2005a **Ab**	Rat (SD); GDs 12−20; 0, 0.1, 1, 10, 30, 100, or 500 mg/kg-d	At 30 and 50 mg/kg-d, disruption of Sertoli-germ cell contact. At 50 mg/kg-d, Sertoli cell hypertrophy, decreased total cell number and number of seminiferous tubules. At 100 mg/kg-d, increased multinucleated gonocytes.	10	30
Kleymenova et al., 2005b	Rat (SD); GDs 12−21; 0 or 500 mg/kg-d	Cytoplasmic changes in Sertoli cells with abnormal cell-cell contact with gonocytes, clustering of gonocytes in the middle of the tubules, altered morphometry of seminiferous tubules, clusters of interstitial cells, decreased number of tubular cross sections per testicular section; increased number of multinucleated gonocytes.		500

Table 4-1. (continued)

Study[a]	Species (strain), duration, and exposure	Reproductive system effects	NOEL mg/kg-d	LOEL mg/kg-d
Lee et al., 2004	Rat (SD); GD 15 to PND 21; 0, 1.5, 14.4, 148, or 712 mg/kg-d (converted from 0, 20, 200, 2,000, and 10,000 ppm DBP in diet)	At 712 mg/kg-d, decreased percent males; decreased AGD and retained nipples, decreased relative testis weight. At 1.5, 14.4, 148, and 712 mg/kg-d, on PND 21, reduction in spermatocyte development, increased foci of aggregated Leydig cells, and decreased epididymal ductular cross section. At 148 and 712 mg/kg-d, at wk 11, loss of germ cell development. At 1.5 mg/kg-d, degeneration and atrophy of mammary gland alveoli in males at 8−11 wks of age.		1.5
Lehmann et al., 2004	Rat (SD); GDs 12−19; 0, 0.1, 1, 10, 30, 50, 100, or 500 mg/kg-d	At ≥50 mg/kg-d, decreased fetal T concentration. At 500 mg/kg-d, a reduction in oil red O staining of lipids in fetal testes.	30	50
Liu et al., 2005	Rat (SD); GDs 12−19; 0, 500 mg/kg-d	Significant reduction in AGD at GD 19.		500
Mahood et al., 2005	Rat (Wistar); GDs 13.5−20.5; 0 or 500 mg/kg-d	Aggregation of fetal Leydig cells; reduced Leydig cell size; reduced T levels at GDs 19.5 and 21.5 (early event in testicular dysgenesis); cryptorchidism; partial absence of epididymis at PND 90.		500
Mylchreest et al., 1998	Rat (SD); GD 30 to PND 20; 0, 250, 500, or 750 mg/kg-d	At 500 and 750 mg/kg-d, decreased AGD. At 250, 500, and 750 mg/kg-d, absent or underdeveloped epididymis, associated with testicular atrophy and germ cell loss, hypospadias, ectopic or absent testes. At 500 and 750 mg/kg-d, absent prostate and seminal vesicles, small testes, and seminal vesicles.		250
Mylchreest et al., 1999	Rat (SD); GDs 12−21; 0, 100, 250, or 500 mg/kg-d	At 500 mg/kg-d, hypospadias; cryptorchidism; agenesis of the prostate, epididymis, and vas deferens; degeneration of the seminiferous epithelium; interstitial cell hyperplasia and adenoma; decreased weight of prostate, seminal vesicles, epididymis, and testes. At 250 and 500 mg/kg-d, retained areolae or thoracic nipples, decreased AGD. At 100 mg/kg-d, delayed preputial separation (attributed to highly affected litter, and not repeated in subsequent study).	100	250

Table 4-1. (continued)

Study[a]	Species (strain), duration, and exposure	Reproductive system effects	NOEL mg/kg-d	LOEL mg/kg-d
Mylchreest et al., 2000	Rat (SD); GDs 12−21; 0, 0.5, 5, 50, 100, or 500 mg/kg-d	At 500 mg/kg-d, decreased AGD, hypospadias, cryptorchidism, absent or partially developed epididymis, vas deferens, seminal vesicles, and ventral prostate; decreased weights of testes, epididymis, dorsolateral and ventral prostates, seminal vesicles, and levator anibulbocavernosus muscle; seminiferous tubule degeneration, focal Leydig cell hyperplasia, and Leydig cell adenoma. At 100 and 500 mg/kg-d, retained thoracic areolae or nipples in male pups.	50	100
Mylchreest et al., 2002	Rat (SD); GDs 12−21; 0 or 500 mg/kg-d	In GDs 18 and 21 fetuses, testicular atrophy, Leydig cell hyperplasia, enlarged seminiferous cords with multinucleated gonocytes; decreased testicular T; fewer epididymal ducts.		500
NTP, 1991	Rat (SD); continuous breeding (16 wks) (gestation and lactation); 0, 80, 385, or 794 mg/kg-d in dams (converted from 0.1, 0.5, and 1.0 % DBP in feed)	F1 adults: At 80, 385, and 794 mg/kg-d: Increased incidence of absent, poorly developed, or atrophic testis and underdeveloped or absent epididymis. At 385 and 794 mg/kg-d: Increased incidence of seminiferous tubule degeneration. At 794 mg/k-d: Decreased mating, pregnancy, and fertility indices; decreased epididymal, prostate, seminal vesicle and testis weights; decreased cauda epididymal sperm concentration; decreased average spermatid count, total spermatid heads/testis or total spermatid heads/g testis; increased incidence of absent, small/underdeveloped/poorly developed, or atrophic penis, seminal vesicles, epididymis, and prostate; interstitial/Leydig cell hyperplasia; delayed testicular descent or cryptorchidism.		80
NTP, 1995 (some of this is also reported in Wine et al., 1997)	Rat (SD); continuous breeding (16 wks) (gestation and lactation); 0, 80, 385, or 794 mg/kg-d in dams (converted from 0.1, 0.5, and 1.0% DBP in feed)	At 794 mg/k-d: Decreased mating, pregnancy, and fertility indices; decreased epididymal, prostate, seminal vesicle, and testis weights.	385	794

Table 4-1. (continued)

Study[a]	Species (strain), duration, and exposure	Reproductive system effects	NOEL mg/kg-d	LOEL mg/kg-d
NTP, 1995	Rat (Fischer 344); perinatal and lactation plus 17 wks; 0, 138, 279, 571, 1,262, or 2,495 mg/kg-d in dam (converted from 0 or 10,000 ppm during gestation and lactation; 0, 1,250, 2,500, 5,000, 7500, 10,000, 20,000 ppm for 4 wks PN; 0, 2,500, 5,000,10,000, 20,000, and 40,000 for last 13 wks PN)	At 571, 1,262, and 2,495 mg/kg-d: Degeneration of germinal epithelium. At 1,262 and 2,495 mg/kg-d: Decreased testes and epididymal weights, fewer sperm heads per testis, and decreased epididymal sperm concentration.	279	571[c]
NTP, 1995	Rat (Fischer 344); perinatal and lactation plus 4 wks; 0, 143, 284, 579, 879, or 1,115 mg/kg-d in dam (converted from 0, 1,250, 2,500, 5,000, 7,500, 10,000, and 20,000 ppm)	At 879 and 1,115 mg/kg-d: Moderate epididymal hypospermia in all males. At 579 mg/kg-d, mild epididymal hypospermia in 2 of 10 males.	284	579[d]
Plummer et al., 2005 **Ab**	Rat (strain not specified); gestation; 0 or 500 mg/kg-d	Decreased fetal T levels.		500
Shultz et al., 2001	Rat (SD), GDs 12−21; 0 or 500 mg/kg-d	Decreased fetal testicular T and androstenedione; increased progesterone.		500
Thompson et al., 2004a	Rat (SD); GDs 12−17, 12−18, or 12−19; 0 or 500 mg/kg-d	Decreased fetal T.		500
Thompson et al., 2005	Rat (SD); GD 19; 0 or 500 mg/kg-d	Decreased fetal T.		500
Wilson et al., 2004	Rat (SD); GDs 14−18; 0 or 1,000 mg/kg-d	Decreased fetal T, expression of *Insl3*.		1,000

Table 4-1. (continued)

Study[a]	Species (strain), duration, and exposure	Reproductive system effects	NOEL mg/kg-d	LOEL mg/kg-d
Zhang et al., 2004	Rat (SD); GD 1 to PND 21; 0, 50, 250, or 500 mg/kg-d	At 250 and 500 mg/kg-d, decreased AGD; underdeveloped epididymides; decreased epididymis or prostate weight at PND 70; decreased percent motile sperm and total sperm heads; degeneration of the seminiferous epithelium. At 500 mg/kg-d, cryptorchidism, absent epididymides, decreased total number of sperm.	50	250

Ab, Abstract only; AGD, anogenital distance; GD, gestation day; PND, postnatal day; LOEL, lowest-observed-effect level for male reproductive system outcomes found in the study; NOEL, no-observed-effect level for male reproductive system outcomes; SD, Sprague-Dawley; T, testosterone. Note: These terms are used solely in a descriptive manner in this table, they may not reflect the terminology of the source study, and they are not intended to convey any regulatory implication.

[a]All studies used an oral route of exposure. Lee et al. (2004) and NTP (1995, 1991) administered DBP in the diet. All other studies used oral gavage.
[b]The abstract states that the effects were "dose dependent," but a LOEL is not specifically indicated.
[c]Overall, the study NOEL and LOEL are lower based on liver peroxisome activity.
[d]Overall, the study NOEL and LOEL are lower based on increased liver weight.

It is also noted that although BMDL values were calculated for specific developmental endpoints identified by Lehmann et al. (2004), Mylchreest et al. (2000), and the NTP (1995) (see Table 4-4 of the 2006 external peer review draft Tox Review for DBP), these values were not used as a POD for reference value derivation.

4.2. REVIEW OF THE TOXICOLOGY DATA SET

Figure 4-1 illustrates the stepwise approach taken in the evaluation of the toxicity studies, focusing on low-dose and low-incidence outcomes. First, for each toxicology study, we examined the data at the lowest dose levels (as defined by the study NOELs and LOELs) (Step 2). If there was any indication of insurmountable problems with the quality of the reported data (e.g., excessive variability, critical methodological concerns, lack of peer review as with abstracts, etc.), or if there were no individual animal data reported (as is often the case for poster abstracts as well as for many published studies which only contain extracted summary data), the review of that study would be terminated. However, if individual data were available, the review could proceed (Step 3). The individual animal data were examined for evidence of reproductive system outcomes in the males. Although for most studies the exposures were only administered during the perinatal developmental period, we recognized that an adverse treatment-related outcome might be identified at any life stage that was assessed in the study. There were three possible courses that the data review could take from this point forward. In cases where problems were identified in the data, we attempted to analyze the extent of the issues and determine the ability to move forward with the study analysis. In some cases, the analysis stopped at this point, due to deficits in the study data or to inadequate reporting of individual animal data. However, if the data in the report appeared to be thoroughly assessed, then the study outcomes and endpoints were examined. Alternatively, in some cases where adequate individual study data were available for analysis (NTP, 1991, 1995), further examination of the study could identify effects at the lowest dose levels that had been considered biologically irrelevant in the original review, but it might require further consideration. At any point in this stepwise process where data were deemed insufficient to proceed further, we identified research needs (discussed in Chapter 7).

To begin the characterization and evaluation of the published studies according to this stepwise model, important aspects of each study protocol, conduct, and reporting were

summarized (see Table 4-2). Examination of this table demonstrates that approximately half of the studies that were selected for analysis (i.e., 14 of 29) were limited to a single dose group, which eliminated them from further examination for lower-dose level effects. It is also important to note that individual animal data were reported in only 2 of the 29 studies, thereby severely limiting, and in some cases even preventing, more rigorous evaluation of the study findings. These two characteristics alone tend to overshadow any of the other listed study attributes that might contribute to confidence in study findings (i.e., evidence that the study was conducted according to quality laboratory standards, description of statistical analysis of the data, and/or specific information regarding the number of litters and offspring assessed, which would provide an indicator of statistical power). Of the studies listed, only those conducted by NTP (1995, 1991) were considered suitable for extended examination.

In order to create a profile of outcomes to the male reproductive system following developmental exposures, which might then serve as a baseline for further comparison and analysis of toxicological findings across the studies, a list of observed effects was compiled (see Table 4-3). The content of this list is very clearly defined by the study protocols, both in terms of what endpoints were examined in each study and when (i.e., at what life stage) they were examined. For some endpoints, the precise GD or postnatal day (PND) of evaluation may even be critical. For example, fetal T should peak at approximately GD 18, so assessments made at earlier or later time points may be less sensitive in detecting adverse outcomes, and the effects will not be directly comparable across fetal ages. Decreases in T levels may not be observed postnatally unless treatment is continued or if testicular malformations disrupt T level (which is a different mechanism of perturbation than alterations to the steroidogenesis pathway). In neonates, examination for nipple retention is generally conducted at around PND 13 when the structure is readily visible but before it is obscured by hair growth. Cryptorchidism, even though present at birth, may not be readily observable in neonates until they reach the age of PND 16−21 (and of course, it should be detectable at postweaning ages and in adults). Preputial separation (PPS) delays can only be observed at the time of sexual maturation, which, in the male Sprague-Dawley (SD) rat, occurs at approximately PND 42; therefore, this effect cannot be detected at an earlier life stage, nor will it be observed in sexually mature adults. On the other hand, sperm alterations (count, morphology, or motility) and perturbations in male fertility can only be assessed in adult males, not in immature individuals at earlier life stages

Table 4-2. Reporting and study size characteristics of male reproductive studies following *in utero* exposure to DBP

Study	> One high dose	Individual data publicly available	Statistical analysis method reported	Study conduct level reported	Number evaluated/group	
					Litters	Offspring
Barlow and Foster, 2003			✓	✓	1−9[a]	7−60[a]
Barlow et al., 2003		✓subset[b]	✓	✓	NR	3
Barlow et al., 2004	✓		✓	✓	8−11[a]	35−74[a,c]
Bowman et al., 2005			✓	✓	18	All male fetuses
Carruthers and Foster, 2005			✓	✓	1−14[d]	1−91[d]
Ema et al., 1998	✓		✓		11 DBP treated	AGD: NR; crypt.: 144
Ema et al., 2000b	✓		✓		73 DBP treated	~770[e]
Ferrara et al., 2006			✓	✓	"in most instances" ~3−6	1−3/litter
Fisher et al., 2003			✓	✓	NR	Testis wt: 5−10 animals/age group (4); hyp. and crypt.: 10 adults
Gray et al., 1999	✓ PPS only		✓		4 (LE); 8 (SD)	LE: 30 male pups; 13 adult males SD: 48 male pups; 17 adult males[f]
Kim et al., 2004 **Ab**	✓				NR	NR
Kleymenova et al., 2004 **Ab**	✓				NR	NR
Kleymenova et al., 2005a **Ab**	✓				NR	NR
Kleymenova et al., 2005b			✓	✓	3	14−21 pups/evaluation

Table 4-2. (continued)

Study	> One high dose	Individual data publicly available	Stat analysis method reported	Study conduct level reported	Number evaluated/group	
					Litters	Offspring
Lee et al., 2004	✓		✓		6−8	11−20 adults
Lehmann et al., 2004	✓		✓	✓	1−4	3−4 fetuses/group
Liu et al., 2005			✓	✓	3	3 fetuses/litter
Mahood et al., 2005			✓	✓	2−7	NR
Mylchreest et al., 1998	✓		✓	✓	7−10	All males/litter
Mylchreest et al., 1999	✓		✓	✓	10	All males/litter
Mylchreest et al., 2000	✓		✓	✓	11−20	All males/litter
Mylchreest et al., 2002			✓	✓	5−6	23−49 fetuses
NTP, 1995, 1991	✓	✓	✓	✓	20	All pups/litter in-life thru necropsy; histopath: 10/selected groups
Plummer et al., 2005 **Ab**					NR	NR
Shultz et al., 2001			✓	✓	3	1 male/litter
Thompson et al., 2004a			✓	✓	4	1 male/litter
Thompson et al., 2005			✓	✓	4	3 fetuses/litter
Wilson et al., 2004			✓	✓	3	All males/litter
Zhang et al., 2004	✓		✓		14−16	20 pups/group

Table 4-2. (continued)

Ab, Abstract only; LE, Long Evans; NR, Not reported; PPS, preputial separation; ✓, present.

[a]Litters and pup numbers not reported for AGD and areolae retention.

[b]Data for three individual animals were reported for LC and Sertoli cell staining. The other results are not reported in this table because they were from toxicogenomic studies (see Chapter 5).

[c]57−100% of these pups survived to necropsy so for malformations that required necropsy, the number of pups is less than shown.

[d]Litters for AGD were the statistical unit; neither litter nor pup numbers for AGD were reported.

[e]Number derived from the mean number of live fetuses/litter.

[f]In some cases, data from two experiments were combined.

Table 4-3. Life stage at observation for various male reproductive system outcomes assessed in studies of developmental exposure to DBP

Findings	Life stage of animals (rats) at observation		
	Fetus	Neonate through puberty	Adult
Decreased T	✓	✓	✓
Malformations	✓	✓	✓
Decreased AGD		✓	✓
Hypospadias		✓	✓
Retained nipples/areolae		✓	✓
Cryptorchidism		✓	✓
Delayed PPS		✓	
Organ weights		✓	✓
Histopathology of male reproductive organs	✓	✓	✓
Abnormal sperm			✓
Decreased fertility			✓

T, Testosterone; AGD, Anogenital distance; PPS, Preputial separation.

Using the list in Table 4-3 as a guide, a more extended analysis was conducted for each of the selected studies. Table 4-4 presents the detailed results. In this table, the various observed outcomes are arrayed across three general life stage categories: prenatal (i.e., observations conducted in fetuses), neonatal through puberty (i.e., observations conducted in pups), and adult (i.e., observations conducted in young, sexually mature animals). These life stage categories do not represent the period of exposure for the study. While all studies include exposures during late gestation (i.e., during the critical window of male reproductive system development), some studies also maintained exposures during later life stages. For reference, Table 4-1 provides general descriptions of exposure durations.

Table 4-4. Age of assessment for individual endpoints across studies of the male reproductive system following developmental exposure to DBP

	Fetus		Neonate through puberty								Adult									
	↓ T[a]	Histo-path[b]	↓ AGD	Hyp	Ret. nip/ areolae	Crypt[c]	Del. PPS[d]	↓ Org wt	Histo-path[b]	↓ T[a]	Malf	↓ Org wt	Histo-path[b]	Ab. Sperm	↓ Fert	Hyp	Ret. nip/ areolae	Crypt	Δ AGD	↓ T[a]
Barlow and Foster, 2003		✓	✓	✓	✓	✓	✓—		✓		✓	✓[e]	✓	✓		✓		✓		
Barlow et al., 2003		✓																		
Barlow et al., 2004			✓		✓						✓	✓[f]	✓			✓[g]	✓	✓	✓↓	
Bowman et al., 2005		✓[h]																		
Carruthers and Foster, 2005			✓[i]		✓[j]						✓	✓	✓			—	✓[k]	—	✓↑	
Ema et al., 1998		✓	✓[l]			✓[l]														
Ema et al., 2000b			✓[l]			✓[l]														
Fisher et al., 2003	✓	✓						✓		✓/—[m]	✓	✓	✓	✓	✓	✓		✓		—[m]
Gray et al., 1999b			✓		✓		✓[n]				✓	✓	✓	✓[o]	✓[p]	✓	✓	✓		✓[p]
Kim et al., 2004 **Ab**				✓		✓		✓		✓[q]										
Kleymenova et al., 2004 **Ab**		✓																		
Kleymenova et al., 2005a **Ab**		✓																		
Kleymenova et al., 2005b		✓		—		—			✓											
Lee et al., 2004			✓	—[r]	✓	—[r]	—	—	✓		—[r]	✓	✓							
Lehmann et al., 2004	✓																			
Liu et al., 2005			✓[s]																	
Mahood et al., 2005	✓	✓				✓[t]		NR	✓		✓	NR	✓			NR		✓		
Mylchreest et al., 1998			✓	✓		✓	—				✓	✓		—[u]		✓		✓		

Table 4-4. (continued)

	Fetus		Neonate through puberty								Adult									
	↓ T[a]	Histo-path[b]	↓ AGD	Hyp	Ret. nip/ areolae	Crypt[c]	Del. PPS[d]	↓ Org wt	Histo-path[b]	↓ T[a]	Malf	↓ Org wt	Histo-path[b]	Ab. Sperm	↓ Fert	Hyp	Ret. nip/ areolae	Crypt	Δ AGD	↓ T[a]
Mylchreest et al., 1999			✓	✓	✓	✓	✓				✓	✓	✓			✓		✓		
Mylchreest et al., 2000			✓	✓	✓	✓	—				✓	✓	✓			✓		NR		
Mylchreest et al., 2002	✓	✓	NR[s]																	
NTP, 1991				✓							✓	✓	✓	✓	✓	✓		✓		
Plummer et al., 2005	✓					✓[v]														
Shultz et al., 2001	✓		NR[s]																	
Thompson et al., 2004a	✓																			
Thompson et al., 2005	✓																			
Wilson et al., 2004	✓																			
Zhang et al., 2004			✓	—		✓		—			✓	✓	✓	✓[u]		—		✓		

✓ Observed; —, Not observed; white box, Not evaluated; shaded box, Evaluated; NR, Not reported, although the study indicates that the endpoint was evaluated; **Ab**, Abstract only; PPS, preputial separation; AGD, anogenital distance; Hyp, hypospadias; Δ, change; Ret. nip/areolae, retained nipples and/or areolae; Cryp, cryptorchidism; Del, delay; Org wt, organ weight decrease (absolute or relative) in at least one reproductive organ; Malf, malformations including ventral/dorsal/lateral prostate, seminal vesicles, androgen dependent muscles, (accessory sex organs) epididymis, vas deferens external genitalia, cryptorchidism, small or flaccid testes; Fert, fertility; Ab. Sperm, abnormal changes in sperm count, motility, and/or morphology.

[a]Decreased testicular testosterone (T) was measured in the fetus; Decreased serum T was measured postnatally and in adults.
[b]Histological changes—Leydig cell hyperplasia (aggregation); multinucleated gonocytes; Wolffian duct increased coiling (can be measured in fetus, neonate through puberty, or adult).
[c]Cryptorchidism can be observed between PNDs 16−21 and older.
[d]Delayed preputial separation normally observed ~PND 42.
[e]Enlargement of the seminiferous cords was observed at PNDs 19−21.
[f]In addition to the observed decreases and absences of male reproductive organs, “occasional enlargement” of the testes was observed only in the 500 mg/kg-d group.

Table 4-4. (continued)

[g]Assessed in adult animals at PNDs 180, 370, and 540. Hypospadias only observed in the 500-mg/kg-d group.
[h]Wolffian ducts smaller, more fragile, adipose tissue surrounding duct was more gelatinous, and decreased coiling.
[i]Assessed at PNDs 1 and 13. Reduction in AGD observed in animals exposed to DBP on GDs 16 and 17, GDs 17 and 18, or GDs 19 and 20; no change in AGD in animals exposed GDs 14 and 15.
[j]Assessed on PND 13; assessed on an individual animal basis, significant increase in nipple retention was observed after dosing on GDs 15−16; 16−17; 17−18; or 19−20.
[k]Assessed at PND 90; significant increase in nipple retention only for males dosed GDs 16−17 (individual animal basis).
[l]AGD and cryptorchidism were assessed in fetuses on GD 21. Exposed pregnant dams were sacrificed on day 21, and live fetuses were removed.
[m] Blood plasma T levels significantly reduced on PND 25 but not on PNDs 4, 10, or in adult.
[n] Delayed PPS only reported for parental generation (P0) males exposed from weaning through to puberty.
[o] Reduced epididymal sperm numbers; not necessarily abnormal sperm.
[p] In P0 males.
[q] Evaluated T levels at 31 and 42 days (not fetus) and found decreased at 42 days.
[r]It is presumed that specific malformations would have been observed if present based on the study design and methods.
[s]Examined in GDs 19 or 21 fetuses.
[t] Observed at PNDs 25 and 90; nonscrotal testes were not evaluated histopathologically.
[u] Only motility was evaluated in Mylchreest et al. (1998); in Zhang et al. (2004), sperm number, motility, and morphology were evaluated, but only count was affected.
[v] Study mentions that adult cryptorchidism was observed, but study methods do not indicate that offspring were retained until adult age.

Table 4-4 summarizes the outcomes and presents a broad representation of positive and negative observations in a manner that demonstrates that not all relevant endpoints were evaluated at all life stages or even in each study. To facilitate summarization of the myriad individual study findings, information was often combined by category (e.g., "histopathology" includes a broad variety of outcomes in various reproductive organs), and for the sake of brevity, the minute details and nuances of the study design and observations, although quite interesting, are not typically presented. In a few cases, negative outcomes presented in the table are extrapolations based on the presumption that specific findings would have been observed if they were present. For example, with methods that include detailed external and internal (macropathology) examination of pups and/or adults, the absence of reported malformations at either of these life stages was presumed to indicate that no gross malformations were observed because they should have been readily detectable (Lee et al., 2004).

Tables 4-1, 4-3, and 4-4 clearly illustrate that the study protocols varied quite extensively. In general, with the exception of the NTP studies, the protocols were not designed to conform to a particular regulatory guideline. Rather, the majority of the studies were focused research efforts that were verifying and/or expanding upon previously observed outcomes; therefore, the differences across study methods are understandable. As a result, the apparent lack of consistency in male reproductive system observations across studies is generally attributable to differences in protocol design and implementation. Some examples are discussed in detail as follows:

- Although all of these studies used exposures during late gestation (i.e., a critical period of male reproductive system development in the rat), the specific endpoints that were assessed and/or the life stages at which endpoints were examined varied extensively across the studies. Obviously, treatment-related alterations of life-stage-specific events require examination during the most appropriate or optimal life stage (e.g., increased multinucleated gonocytes can only be observed in fetal testes, delays in PPS can only be observed in juvenile animals at the time of sexual maturation, and disturbances in reproductive function can only be observed in sexually mature adults). Other permanent structural abnormalities may be detected across multiple life stages (e.g., hypospadias or cryptorchidism could theoretically be observed in late gestation fetuses, in adolescents, and in adults). For some outcomes, it is difficult to predict *a priori* the optimal time point for evaluation. For example, DBP-related increases in the estrogen receptor (ER) were observed at 31 days but not at 42 days (Kim et al., 2004).

- It is important to realize that not all available offspring are evaluated in every study; therefore, identification of adverse outcomes may rely, in part, on sampling protocols and the statistical power of the sample size for detection of rare or low-incidence events. Calculations of statistical power are rarely provided in study reports.

- In some cases, apparent differences in studies may result because the report contains an insufficient level of detail on a particular endpoint or life stage—often because the emphasis of the scientific review lies in a slightly different direction. For example, if high doses of DBP are administered during sensitive periods of male reproductive system development, and the males are maintained and terminated as adults, at which time histopathological evaluation is performed, it might be assumed that various male reproductive system malformations and/or cryptorchidism would have been present in some of the males at necropsy. Yet, these findings may not be reported because the histopathological findings are the primary focus of the investigation and/or the publication (e.g., Lee et al., 2004).

- In other situations, the description of the findings at various life stages may vary. For example, evidence of cryptorchidism may be described as "testis located high in the abdomen" in a fetus, as "undescended testis(es)" in an adolescent rat, or as "unilateral testis" upon noninvasive clinical examination of an adult. To some extent, this lack of consistency in terminology may result from laboratory Standard Operating Procedures that direct technical staff to avoid the use of diagnostic terminology.

Overall, in spite of numerous differences in the study designs, the toxicity data set for DBP clearly demonstrates that exposure to DBP during critical stages of male reproductive system development can result in adverse structural and functional reproductive outcomes. When specific critical aspects of study design and implementation were similar, consistent outcomes were almost universally observed. The WOE embodied by the data described above is further supported by studies in rats that demonstrated similar incidences of cryptorchidism and decreased AGD in male pups of dams treated with either DBP or MBP, the metabolite of DBP (Ema and Miyawaki, 2001a). The ability of MBP to cross the placenta and reach the fetus has also been conclusively demonstrated (Fennell et al., 2004; Saillenfait et al., 1998), and these two TK events (metabolism and placental transport) are key to the MOA of reduced fetal testicular T (David, 2006). Available toxicogenomic data, described elsewhere in this case-study report, further elucidate the MOA(s) of DBP in producing adverse effects on male reproductive system development and are an important consideration in the WOE analysis of the toxicological data set.

In the selected DBP toxicology study data set, the presentation of extensive individual offspring data was limited to the NTP (1991) study conducted as a reproductive assessment by continuous breeding (RACB) in SD rats. The individual data from this study were carefully examined in order to confirm the NOEL and LOEL described in the study report. This analysis was conducted under the presumption that statistical and/or biological significance noted in the summary compilations of male reproductive system outcomes might not identify low incidence effects in individual offspring at lower dose levels. To further aid the identification of treatment-related outcomes, the male reproductive system outcomes were grouped by organ instead of individual animal. This analysis revealed apparently treatment-related findings in the testis and epididymis of F1 male offspring, as summarized in Table 4-5. At the highest dose tested (794 mg/kg-d, equivalent to 1.0% DBP in the diet), additional findings in the male reproductive organs of F1 offspring included single incidences of (1) underdeveloped prepuce; (2) mild secretion and severe vesiculitis of the prostate; (3) a mass on the testis; and (4) a focal granuloma with fluid and cellular degeneration in the epididymis; none of these findings were observed at the lower dose levels. Understandably, the findings at the low- and mid-dose groups were originally interpreted as not being treatment-related (Wine et al., 1997; NTP, 1991). However, consideration of MOA information for DBP, including toxicogenomic data, resulted in a more conservative interpretation of the toxicity data both by NTP researchers (conference call in 2008 between Paul Foster [NTP/NIEHS], Susan Makris [EPA/NCEA], and Susan Euling [EPA/NCEA]) and by the EPA IRIS program (U.S. EPA, 2006a). Consequently, further analysis of individual offspring data in the current case study did not identify any additional sensitive toxicological outcomes; the study LOEL was confirmed to be the lowest treatment level tested in the NTP RACB study (80 mg/kg-d).

4.3. UNEXPLAINED MOAs FOR DBP MALE REPRODUCTIVE TOXICITY OUTCOMES

Figure 3-6 illustrates the broad conceptual approach for consideration and interpretation of toxicogenomic and toxicology data to inform MOA. The toxicogenomic data can be evaluated to identify altered genes, gene products, and pathways; this information can lead to a more complete understanding of the mechanism of action or MOA(s) for the chemical toxicity. From the opposite perspective, the toxicity data can provide information critical to identifying

Table 4-5. Incidence of gross pathology in F1 male reproductive organs in one continuous breeding study with DBP[a]

Gross finding[b]	Dose (% in diet)			
	0	0.1	0.5	1.0
Testis: absent, poorly developed, atrophic, undescended	0/20	1/20	1/20	6/20
Penis: small/underdeveloped	0/20	0/20	0/20	4/20
Epididymis: underdeveloped/absent	0/20	1/20	1/20	12/20

[a]Incidences were compiled from reported individual animal macroscopic pathology data; statistical analysis was not performed.
[b]Some animals have more than one type of malformation, and these animals were counted separately for each of the three outcome categories.

Source: (NTP, 1991).

the relevant MOA(s) involved in the toxicological outcomes, and thereby inform the interpretation of gene alterations and relevant pathways.

Each male reproductive system outcome was evaluated for consistency with either or both of the two well-established MOAs using expert judgment based on the available published literature for DBP (see Figure 4-2). This exercise helped to identify the unexplained endpoints for which the evaluation of the toxicogenomic data set may suggest potential MOAs (see Chapter 5). For the DBP case study, Table 4-6 presents a compendium of the specific findings noted in the male reproductive system following exposures at critical windows of development. While reduced fetal testicular T and reduced *Insl3* signaling can be linked to some of the observed outcomes on the basis of available data, potential key events cannot specifically be identified for other outcomes. The unexplained MOAs are good candidates for further study, both in toxicology and toxicogenomic studies, to elucidate the underlying mechanism of action.

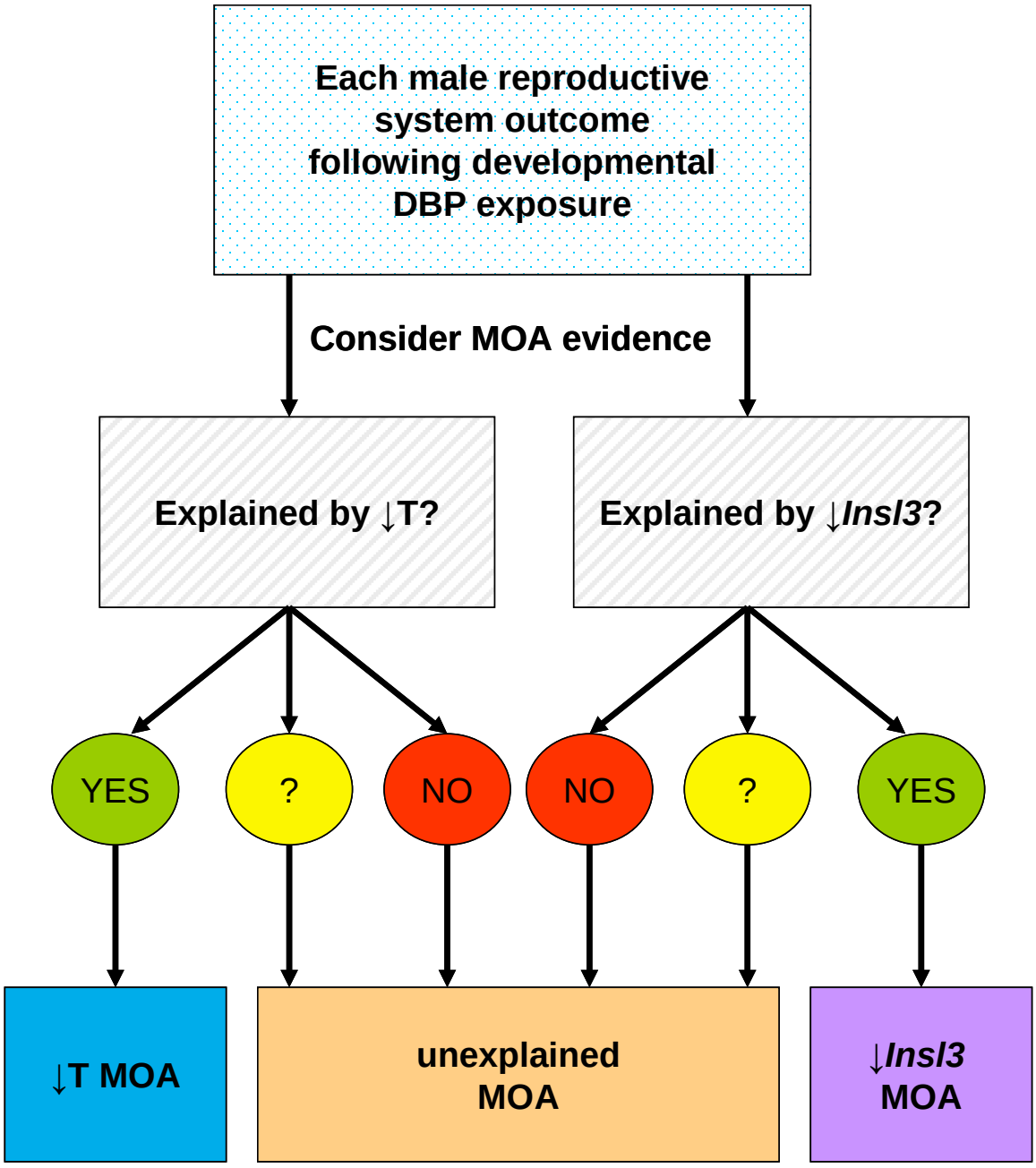

Figure 4-2. The process for evaluating the MOA for individual male reproductive system outcomes following developmental DBP exposure. The available data for MOA for each male reproductive outcome following developmental DBP exposure were evaluated by a team of experts. For each outcome, the current WOE of the data either support the MOA ("YES"), support that this is not the MOA ("NO"), or are inconclusive for the MOA, i.e., either unlikely or unclear ("?"). "Unexplained MOAs" include both "?" and "NO" conclusions.

Table 4-6. Evidence for MOAs for the observed effects in the male reproductive system after *in utero* DBP exposure

Organ/ Function	Effect	MOA	
		Reduced fetal testicular T	Reduced *Insl3* signaling
Testes	Multinucleated gonocytes; increased number of gonocytes in fetal testes	?[b]	?[c]
	Altered proliferation of Sertoli and peritubular cells; fewer Sertoli cells	?[b]	?[c]
	Gonocyte apoptosis increase; early postnatal decrease in gonocyte number	?[b]	?[c]
	Abnormal Sertoli cell-gonocyte interaction	?[b]	?[c]
	Small incidence of Leydig cell adenomas, aggregates, and hyperplasia	✓	?[c]
	Decreased number of spermatocytes or cauda epididymal sperm concentration.	✓	✓[d]
	Small or flaccid; other abnormalities; decreased weight	✓	✓
	Increased weight due to edema	?[e]	?
	Decreased number or degeneration of seminiferous cords/tubules; altered morphology; degeneration of the epithelium; enlarged cords/tubules	?[b]	?[c]
	Testes descent: none (cryptorchid) or delayed	✓[f]	✓[f]
Gubernacular ligament	Gubernacular ligament development effects: agenesis or elongation	X	✓
Epididymis	Lesions and agenesis; partial to complete absence; decreased epididymal ductular cross section	✓	X
	Reduced weights	✓	✓
Mammary gland	Nipple and/or areolae retention in males	✓	X
	Degeneration and atrophy of alveoli in males	?[b]	X

Table 4-6. (continued)

Organ/ Function	Effect	MOA	
		Reduced fetal testicular T	Reduced *Insl3* signaling
Wolffian ducts	Underdeveloped	✓	X
Seminal vesicles	Malformations or absent; decreased weight	✓	X
Coagulating gland	Malformations	✓	X
Penis	Small, underdeveloped	✓	X
	Hypospadias	✓	X
	Delayed preputial separation	✓	X
Accessory sex organ	Decreased weight	✓	X
Prostate	Decreased wt or absent	✓	X
Vas deferens	Agenesis	✓	X
Levator anibulbocavernosus muscle	Decreased weight	✓	?[c]
Male/female ratio	Decreased % male offspring as determined by AGD at birth	✓	X
Perineum	Decreased AGD	✓	X
Repro function	Infertility	✓	✓[d]

AGD, anogenital distance; ?, Current data indicate that it is unlikely the MOA; ✓, Current weight of evidence of the data support this MOA leading to the effect; X, Current weight of evidence of the data indicate that this MOA is not the MOA for this outcome.

[a]MOA is defined as one or a sequence of key events upon which the outcome is dependent (see glossary).
[b]Reduced fetal testicular T may play a role, but current data indicate that reduced T is not solely responsible for this outcome.
[c]The *Insl3* knockout mouse phenotype suggests that *Insl3* is specifically required for gubernacular ligament development and, therefore, testis descent in mice since these mice do not have other defects.
[d]Decreased fertility in males is a result of reduced *Insl3* signaling since reduced *Insl3* signaling leads to undescended testes, which, in turn, reduces sperm count (presumably by increasing the temperature) and can cause infertility.

Table 4-6. (continued)

[e]In some animals, increased weight, due to edema, can result in animals that have epididymal agenesis, which is a consequence of reduced testosterone (T).

[f]*Insl3* signaling is required for development of the gubernacular ligament and through this mechanism—the 1st stage of testis descent from the kidney region to the inguinal region. Testosterone is required for the 2nd stage of testis descent, from the inguinal region to the scrotum (reviewed in Klonisch et al., 2004). After *in utero* DBP exposure, the cryptorchid phenotype resembles the *Insl3* knockout. A delay in testis descent can result from reduced *Insl3* and T.

4.4. CONCLUSIONS ABOUT THE TOXICITY DATA SET EVALUATION: DECISIONS AND RATIONALE

The review of the toxicology data set identified a number of issues and limitations that are evident in the study descriptions and endpoint summaries presented in this chapter. These include the following:

- *Lack of dose-response information:* A number of studies conducted with DBP used a single high-dose treatment level (often at 500 mg/kg-d) in order to produce readily observable adverse outcomes to male reproductive system development that could be examined. In such studies, the absence of lower dose levels prevents the evaluation of dose-dependent responses and does not allow the identification of study-specific NOELs or LOELs. While this approach is useful for hazard characterization, it does not facilitate other aspects of risk assessment (e.g., dose-response assessment or risk characterization). Thus, studies utilizing a single high-dose level may provide important information for a WOE assessment of the toxicology profile, but they have diminished usefulness in identifying outcomes for use in risk calculations at environmentally relevant doses.

- *Insufficient information on study methods:* Even though every study report includes a section on study methods, there can be a great deal of unevenness in the amount of detailed information provided. Consequently, important questions may arise during study review that cannot be readily resolved. In some cases, this can have an impact on individual study interpretation or on conclusions that rely on a thorough WOE evaluation of the data set.

- *Unavailable individual outcome data*: A full range of individual animal data is seldom included in studies published in the open literature and is almost never available when the only available publication is a presentation abstract. Conversely, individual animal data are generally included in toxicology reports generated in response to a regulatory mandate or conducted by a federal agency (e.g., NTP). The availability of individual animal data can be quite important in interpreting the study findings, because it can reveal problems or inadequacies in the data, but it can also help identify low incidence adverse outcomes. In the case of DBP, the individual offspring data presented in the NTP study report (1991) include alterations in the reproductive system of F1 males that

had been exposed during development. These findings are similar to outcomes identified at higher dose levels, are consistent with the proposed MOA, and, consequently, are used to establish a LOEL for the study.

- *Protocol limitations:* Unless studies are designed to meet the recommendations of a standardized testing protocol (e.g., NTP or U.S. EPA/Office of Prevention, Pesticides and Toxic Substances reproductive toxicity study guidelines), there may be a high degree of variability among the protocols used for testing any one chemical. Between two studies, there can be differences in the treatment regimen or in the assessment of outcomes that render them incomparable. DBP provides a good example of a chemical that targets a very specific critical prenatal window of reproductive system development in males, and results in adverse outcomes that could go unidentified if the appropriate endpoint(s) are not assessed at the optimal life stage or time point.

- *Specific study's limitations:* Even when a study design optimizes the detection of adverse outcomes from chemical treatment, there may be challenges in study analysis and interpretation. Such is the case with the NTP study on DBP, which was conducted in several phases and reported both in the open literature (Wine et al., 1997) and by the Institute that conducted the experiments (NTP, 1995, 1991).

The analysis of the toxicology data in this chapter has provided a firm basis for expanded consideration of the toxicogenomic data for DBP as depicted in Figure 3-6. The extensive analysis of the toxicology data set and consideration of MOA(s) provide a source of information for use in phenotypic linking of known and potential MOAs. This chapter also provides an example of steps one can take to develop a toxicological data source, in particular, examining (1) the individual toxicity studies; (2) the WOE for the studies; (3) potential low incidence and low-dose effects; and (4) the MOA for the affected endpoints. All of these steps are useful exercises for evaluating toxicogenomic data in future risk assessments. The evaluations of both the toxicity and toxicogenomic data sets (detailed in Chapters 4 and 5) provide strong support for phenotypic anchoring for a number of gene expression changes occurring after *in utero* DBP exposure for several of the male reproductive outcomes. The available toxicogenomic studies for DBP are evaluated in Chapter 5.