別添 4

厚生労働行政推進調査事業費補助金（化学物質リスク研究事業）
OECD プログラムにおいて TG と DA を開発するための AOP に関する研究

令和２年度 分担研究報告書

AOP、TG、DAの開発、AOP国内マニュアルの作成

研究分担者　小島　肇
国立医薬品食品衛生研究所　安全性生物試験研究センター　安全性予測評価部　室長

研究要旨

動物実験 3Rs の国際的な浸透に加えて、実験動物とヒトとの種差等の克服のために、既存の毒性試験法の見直しが世界的に進んでいる。経済協力開発機構（OECD: Organisation for Economic Co-operation and Development）においても、全身毒性試験（免疫毒性、生殖発生毒性、発がん性及び光安全性等）の有害性発現経路（AOP: Adverse Outcome Pathway）を開発し、その情報を活用して動物実験代替法を念頭においた試験法ガイドライン（TG: Test Guideline）の公定化や *in silico* 法を確立する一方で、AOP 等の毒性情報を網羅した“試験の実施と評価のための戦略的統合方式(IATA：Integrated Approaches to Testing and Assessment）”を開発し、それに基づき、TG と同格の扱いになる確定方式（DA：Defined Approach）による化学物質の安全性評価を推進している。このような国際的な潮流に乗り、日本が得意とする分野で主導権を握って、AOP や TG を公定化し、さらには IATA や DA の開発に協力することが本研究班の目的である。

昨年度からの継続した活動の中、日本人の開発した以下の TG 3 件の改定が令和 2(2020)年 6 月に OECD により公表された。

1) AR STTA 法 ： AR-EcoScreen™ 細胞を用いた アンドロゲン受容体恒常発現系転写活性化試験（TG458）
2) 眼刺激性試験　短時間曝露法（TG491）
3) 皮膚感作性試験代替法 アミノ酸誘導体反応試験（ADRA: Amino acid Derivative Reactivity Assay）（TG442）

AOP に関しては、“Inhibition of Calcineurin Activity Leading to impaired T-Cell Dependent Antibody Response”が日本発の AOP の一つとして、令和 2(2020)年 12 月に OECD, EAGMST (Extended Advisory Group on Molecular Screening and Toxicogenomics)において内諾となった。

研究協力者

相場節也
東北大学医学系研究科・医学部
皮膚科学分野教授

足利太可雄
国立医薬品食品衛生研究所
安全性生物試験研究センター
安全性予測評価部 主任研究官

稲若邦文
一般社団法人日本化学工業協会
化学品管理部長
大石　巧
日本免疫毒性学会試験法委員会
AOP 検討小委員会
尾上誠良
静岡県立大学　薬学部・薬剤学分野
教授
笠原利彦
富士フィルム株式会社
安全性評価センター 技術マネージャー
加藤雅一
株式会社ジャパン・ティッシュ・エンジニアリング(J-TEC)主任研究員
木村　裕
東北大学医学系研究科 医学部
皮膚科学分野　助教
久田　茂
日本免疫毒性学会試験法委員会
AOP 検討小委員会

A. 研究目的

本研究班では、OECD(Organisation for Economic Co-operation and Development)のAOP(Adverse Outcome Pathway)開発プロジェクトの中で、化学物質の毒性情報等を集積しながら、免疫毒性、発がん性及び光安全性等に関する日本発の AOP 開発を進める。既存の AOP 情報をもとに開発された皮膚感作性試験代替法 ADRA、免疫毒性試験 MITA(Multi-Immuno Toxicity Assay)、光安全性試験スクリーニング ROS(Reactive Oxygen Species) assay 、 LabCyte EPI-MODEL24 を用いる腐食性試験代替法については、試験法毎に独立した国内外の専門家による第三者評価(peer review または review)を受けた後、TG を開発する。一方で、皮膚感作性DA(DASS: Defined Approach for Skin Sensitisation)の開発に関与することを通じて、IATA(Integrated Approaches to Testing and Assessment)や DA の国内での普及に務める。

B. 研究方法

B.1. AOP、TG、DA の開発、AOP 国内マニュアルの作成

B.1.1. AOP 国内マニュアルの作成

OECD の AOP 開発プロジェクト EAGMST (Extended Advisory Group on Molecular Screening and Toxicogenomics)及び、TG の開発プロジェクト WNT(Working Group of the National Coordinators of the Test Guidelines Programme)の進捗に合わせ、班員を支援した。

この過程で、AOP作成のルール変更に対する日本の貢献をまとめた。

B.1.2. AOP開発

AOP に関しては、足利分担研究者とともに、日本免疫毒性学会会員をメンバーとする同学会試験法委員会 AOP 検討小委員会に免疫毒性 AOP の開発を委託している。

文献調査の結果に基づいて、カルシニューリン阻害を分子初動(MIE: Molecular initiating event)とし、T 細胞依存性抗体産生抑制(TDAR: T-cell dependent antibody response) を有害性発現(AO : Adverse Outcome)とす Inhibition of Calcineurin Activity Leading to impaired T-Cell Dependent Antibody Response（AOP154）を作成し、外部 review に対応した。

B.1.3. TG 開発

AR STTA 法 : AR-EcoScreen^{TM} 細胞を用いた アンドロゲン受容体恒常発現系転写活性化試験 (AR STTA: The Stably Transfected Transactivation method using the

AR-EcoScreen™ cell line)TG458、眼刺激性試験　短時間曝露法 TG491 及び皮膚感作性試験代替法 *In Chemico* Skin Sensitisation,（ADRA: Amino acid Derivative Reactivity Assay）TG442C の改定に向け、尽力した。

B.1.4. OECD 作業計画にある試験法

B.1.4.1. 免疫毒性試験

相場が開発し、他の研究班でバリデーションを終了させた IL-2 を指標とした免疫毒性試験の TG を目指し、海外の専門家を招聘したバリデーション報告書の peer review を実施した。また、*in vitro* 免疫毒性に関する DRP (Detailed Review Paper)を国際的な専門家とともに作成した。

B.1.5. DA の開発協力

足利分担研究者とともに、OECD 専門家会議において、DASS の開発に協力した。

C. 研究結果

C.1. AOP、TG、DA の開発、AOP 国内マニュアルの作成

C.1.1. AOP国内マニュアルの作成

平成 30(2018)年まで本研究班からの提案も含む 8 本の AOP 案を日本から OECD に提案してきた。さらに平成 30(2018)年、本研究班からの提案も含む 17 本の SPSF(Standard Project Submission Form)を提案した。その中で、13 本の「ラットにおける非遺伝毒性発がん性」に関する AOP 案を日本製薬工業協会グループの協力を得て提出した。

これらの過程で AOP 作成のルール変更に対する日本の貢献をまとめた。ただし、目標にしていた国内マニュアルはまだ完成していない。この理由として、国内マニュアルの基になる OECD における AOP ハンドブックが再改定されることになったためである。

貢献の第一は、AOPの行政的な受け入れ委への関与である。医薬品のがん原性評価の多くは､ICHがん原性試験ガイドラインに基づき、げっ歯類がん原性試験により評価されており、げっ歯類で発生が増加した腫瘍について､発癌機序を検討し､ヒトでの発がんリスクを評価する。これまでに医薬品による非遺伝毒性機序による発がんの知見が蓄積されていることから、これらの発がん機序とヒトでのリスク評価に基づいて作成されたAOPは、薬理作用からげっ歯類における発癌とヒトでのリスクを予測するために有用と考えられた。しかしながら､薬理作用等のデータがほとんど無いことが多い一般化学物質の発がん性評価には医薬品とは異なるアプローチが求められる。従って、ヒトのAOP開発を念頭におかず、ラットにおける取り組みを用いたAOPでは、化学物質全般を評価対象とするOECDの枠組みには疑問が呈された。

本件を初動として、以降、EGMAST は WNT や WPHA (Working Party on Hazard Assessment) に AOP の SPSF において、規制との関連性が明記されることになった。具体的には、以下の設問が追加された。

Proposers should indicate if and how the proposed AOPs are associated to any regulatory toxicological endpoints (e.g. acute or chronic toxicity, toxicity to reproduction, developmental neurotoxicity, non-genotoxic carcinogenicity, endocrine disruption etc.). Proposers will indicate what are the potential regulatory applications of the proposed AOPs. The following elements can be considered in addressing this section:

- Is the project linking to ongoing or future projects in OECD such as Integrated Approach to Testing and Assessment (IATA) projects [link to webpage – see case study projects] or Test Guideline development [Link to current OECD TGs - Link to TG development workplan]? (if so, please describe)
- Do the proposed AOPs complement an existing network of AOPs addressing a regulatory endpoint? (if so, please describe)
- Do the proposed AOPs identify a regulatory gap, or lack of adequate testing methods and thus:
  - Help identify candidate in vitro assay or battery of assays (if so, please describe)
  - Help standardise testing for certain endpoints (if so, please describe)

  Proposers should also mention if they are aware of any indications of commitment from any organisation (e.g. government/agency/ academia) to support AOP development and eventual review.

この追加により、SPSF 提出時にナショナルコーディネーターが事前に確認することになった。

もう一点、日本の AOP 開発への貢献が review システムの変更である。AOP 開発のためには、これまで EAGMST における内部 review の後に、外部 review へと進む。AOP に不慣れな日本人対応のため、EAGMST は早期に外部 review に進むべくコーチ制度と導入した。現在、免疫毒性の 3AOP に関しては、コーチとの意見交換による修正を進めている。

C.1.2. AOP開発

Inhibition of Calcineurin Activity Leading to impaired T-Cell Dependent Antibody Response（AOP154）については、外部reviewerからのコメントに対応した結果、外部reviewがほぼ終了した。

令和2（2020）年12月にはEAGMSTの内諾もなされ、WNTとWPHAの了承が得られ次第、OECDの正式なAOPとなる。

C.1.3. TG の開発

AR-STTA　TG458 の改定に向け、追加される 2 試験法 The AR-CALUX® method using the AR-CALUX® cell line 及び The ARTA method using the 22Rv1/MMTV_GR-KO cell line の peer reviewer を務めた。結果として、これら試験法を含む改定 TG458 は令和 2 年 6 月に正式に TG となった(添付資料 1)

眼刺激性試験　短時間曝露法 TG491 に、揮発性物質に関する適用範囲の拡大が議論され、2020 年 4 月に開催された 32nd WNT 会議で改定が採択され、6 月に公表された(添付資料 2)。

ADRA TG442C の改訂については、ADRA の Annex1 Table 1 の習熟度確認物質の中のプロピルパラベンの分子量が 110.1 から 180.2 に修正された(添付資料 3)。引き続き、1)適用濃度を 1mM から 4mM に引き上げにより、偽陰性の改善が期待される。本件については、バリデーション研究を実施することになり、専門家会議での計画審議を受け、令和 2 年 12 月より開発者の富士フィルム主導のもと、5 施設の協力を受け、12 物質を用いるバリデーション研究が別研究班で実施されている。2)混合物を評価するため、蛍光を利用した試験法の追加については、1)の結果と合わせ、来年度に議論される。

3)陽性対照物質を追加する及び 4)性能標準物質を変更するについては、本年の TG442C の改定を前提に議論され、令和 3

（2021）年 4 月の 33rd　WNT 会議で改定案が採択される予定である。

C.1.4. OECD 作業計画にある試験法

C.1.4.1. 免疫毒性試験

表 1 に示す海外の専門家を招聘し、IL-2 Luc assay バリデーション報告書の peer review 報告書が完成した(添付資料 4)。

結論として、バリデーション報告書では、以下のように結論された。

We conducted the validation study of the IL-2 Luc assay among the 4 luciferase assays that comprise the MITA. The results of both Phase I and Phase II studies satisfied the acceptance criteria for the validation study. Although the predictivity could not reach 80%, it may be acceptable when considering its applicability domain and limited target. So, we would like to propose the IL-2 Luc assay for the OECD test guideline of *in vitro* immunotoxicity test.

これを受けた Peer review 報告書では、以下のように結論された。

The PRP concluded that, even though the predictive capacity was not sufficient to allow use as a stand-alone test, the IL-2 Luc assay validation has demonstrated that the method should be acceptable as a part of IATA for the predictive screening of T-cell targeted immunotoxicity

一方、*in vitro* 免疫毒性試験は未だに OECD で採択されたことはない。このような場合、DRP を作成し、その分野の現状を報告することになっている。

そこで、DRP の開発を OECD に提案し、表 2 に示す国際的な専門家の協力を受けて作成した。この DRP を OECD に令和 2(2020)年秋に送ったところ、各国からの意見が寄せられた。この意見をもとに改訂を続けている。

C.1.5. DA の開発協力

OECD 専門家会議（電話会議）で DASS の開発に寄与した。2 週間に一度の電話会議で、データベースの見直しを行うとともに、DASS 案をもとに、試験法の組み合わせにおける予測性、適用限界、不確実性に関する議論が進んだ。

この DASS 案について、令和 3 年 4 月の 33rd　WNT 会議で採択される予定である。

D. 考察

国際的な潮流に乗り、日本が得意とする分野で主導権を握って、AOP や TG を公定化し、さらには IATA や DA の開発に協力することを目指してこの 3 年間取り組んできたが、TG はともかくとしても、AOP や IATA など思うように進んでいない。担当者としてももどかしい毎日を送っている。他国が作成した文書に日本の意見を送るというレベルと日本が中心となって国際的な論文を作成するというレベルまでならともかく、日本が中心となって国際的な規制をまとめることは異次元のものである。世界の中心でいるためには、国内の専門家をまとめ、国際的な合意を計らねばならない。免疫毒性や生殖毒性など初めての取り組む毒性分野ではましてや、抵抗も大きく、想定内とはいかない労苦を伴う。積極的な取り組みをする専門家が孤立しないよう引き続き、怠りない支援を続けていきたい。

E. 結論

昨年度からの継続した活動の中、本年度に TG に関する 3 件の改定が令和 2 年 6 月に OECD により公表された。

1) AR STTA 法 ： AR-EcoScreen™ 細胞を用いた アンドロゲン受容体恒常発現系転写活性化試験（TG458）

2) 眼刺激性試験　短時間曝露法（TG491）
3)皮膚感作性試験代替法 アミノ酸誘導体反応試験（ADRA）（TG442）

AOP に関しては、Inhibition of Calcineurin Activity Leading to impaired T-Cell Dependent Antibody Response（AOP154）が日本発の AOP の一つとして、令和 2 年 12 月に EAGMST により内諾となった。

引き続き、OECD の活動の中で、日本が得意とする分野で主導権を握って、AOP や TG を公定化し、さらには IATA や DA の開発に協力していく予定である。

F. 添付資料

1. OECD Test Guideline 458: Stably Transfected Human Androgen Receptor Transcriptional Activation Assay for Detection of Androgenic Agonist and Antagonist Activity of Chemicals
2. OECD Test Guideline 491: Short Time Exposure In Vitro Test Method for Identifying i) Chemicals Inducing Serious Eye Damage and ii) Chemicals Not Requiring Classification for Eye Irritation or Serious Eye Damage
3. OECD Test Guideline 442C for the Testing Chemicals on in chemico skin sensitisation assays addressing the Adverse Outcome Pathway Key Event on covalent binding to proteins
4. IL-2 Luciferase (IL-2 Luc) Assay Report of the Peer Review Panel

G. 研究発表

G-1. 学会誌・雑誌等における論文一覧

1) 山田 隆志, 足利 太可雄, 小島 肇, 広瀬 明彦：AOP (Adverse Outcome Pathway；有害性発現経路) に基づいた化学物質の安全性評価へ向けたチャレンジ. Yakugaku Zasshi. 2020;140(4): 481-484.doi: 10.1248/yakushi.19-00190-1
2) 小島 肇:OECD 試験法ガイドライン開発における CERI の国際貢献. CERI NEWS, 2020;90:2-3.
3) 小島 肇:AOP 及び IATA に基づく安全性評価手法の進捗. JETOC 40 周年記念誌, 2020;71-101.
4) Mizoi K, Arakawa H, Yano K, Koyama S, Kojima H, Ogihara T: Utility of Three-Dimensional Cultures of Primary Human Hepatocytes (Spheroids) as Pharmacokinetic Models. JBiomedicines. 2020;8(10):374. doi: 10.3390/biomedicines8100374
5) Kato Y, Yamamoto N, Hiramatsu N, Sato A, Kojima H: Inhouse Fabrication of a Reconstructed Human Corneal Epithelium Model for Use in Testing for Eye Irritation Potential. Applied in Vitro Toxicology, 2020;6(3), doi: 10.1089/aivt.2020.0003
6) 尾上誠良, 上月裕一, 豊田明美, 笛木修, 細井一弘, 小島 肇, 足利太可雄, 小野寺 博志:光安全性評価の現状と課題. YAKUGAKU ZASSHI, 2021, 141(1), 111-124. https://doi.org/10.1248/yakushi.20-00148
7) Imamura M, Wanibuchi S, Yamamoto Y, Kojima H, Ono A, Kasahara T, Fujita M: Improving predictive capacity of the Amino acid Derivative Reactivity Assay test method for skin sensitization potential with an optimal molar concentration of test chemical solution, J Appl Toxicol.2021;41(2):303-329. doi: 10.1002/jat.4082.

G.2 学会発表

1) 小島 肇: 安全性評価における Replacement の概要, 第 47 回日本毒性学会学術年会, 2020/6/29, 国内（web 開催）
2) 小島 肇: ICH における発生毒性代替法の考え方, 第 60 回日本先天異常学会学術集会, 2020/7/12, 国内（web 開催）
3) 小島 肇: 化粧品安全性研究はどこまで進んでいるか -国際情報・代替法-, アレルギー成分確認方法のエキスパートセミナー／多職種ワークショップ, 2020/8/30, 国内（web 開催）
4) 深井悠貴, 溝井健太, 松本映子, 小山智志, 矢野健太郎, 石田誠一, 小島 肇, 荻原 琢男: OECD/TG の cytochrome P450 誘導試験における mRNA 測定の有用性, 第 27 回 HAB 研究機構学術年会, 2020/9/3-4, 国内（web 開催）
5) Kojima H, Ishida S: Challenge of standardization in the AMED-MPS project, Global Summit on Regulatory Science (GSRS20), 2020/9/30, 国外（web 開催）
6) 小島 肇: OECD におけるコンピューターモデルの行政的な受け入れ, CBI 学会 2020 年大会, 2020/10/28, 国内（web 開催）
7) 小島 肇: 医薬品の発生毒性試験にゼブラフィッシュ試験を利用する上での期待, 問題点, 課題, 第 6 回ゼブラフィッシュ創薬研究会, 2020/10/30, 国内（web 開催）
8) 赤木隆美, 村上将登, 宮崎裕美, 田口浩之, 池田英史, 加藤雅一, 山田知美, Mura S, Couvreur P, 足利太可雄, 小島 肇, 明石 満:三次元培養皮膚モデル LbL-3D Skin を用いた皮膚刺激性試験代替法の バリデーション研究, 日本動物実験代替法学会 第 33 回大会, 2020/11/12, 国内（web 開催）
9) 水町秀之, 渡辺美香, 生悦住茉友, 梶原三智香, 安田美智代, 水野 誠, 今井教安, 佐久間めぐみ, 芝田桃子, 渡辺真一, 上野順子, Basketter D, Eskes C, Hoffmann S, Lehmann D, 足利太可雄, 寒水孝司, 武吉正博, 宮澤正明, 小島 肇:皮膚感作性試験代替法 Epidermal Sensitization Assay (EpiSensA) の Validation 研究(施設内再現性 Phase I), 日本動物実験代替法学会 第 33 回大会, 2020/11/12, 国内（web 開催）
10) 木村 裕, 安野理恵, 渡辺美香, 小林美和子, 岩城知子, 藤村千鶴, 近江谷克裕, 山影康次, 中島芳浩, 真下奈々, 岡山昂祐, 高木佑実, 大森 崇, 小島 肇, 相 場 節 也 :Multi-ImmunoTox Assay (MITA) : IL-1 Luc assay バリデーション試験の結果, 日本動物実験代替法学会第 33 回大会, 2020/11/12, 国内（web 開催）
11) 山口宏之, 押方 歩, 綿谷弘勝, 小島 肇, 竹澤俊明:固体を含む被検物質に使用可能な Vitrigel-EIT 法の改訂手順の提案, 日本動物実験代替法学会 第 33 回大会, 2020/11/12, 国内（web 開催）
12) 浅野哲秀, 笠松俊夫, 北本幸子, 山本美佳, 足利太可雄, 小島 肇: Bhas42 細胞形質転換試験法（Bhas 42 CTA）の評価, 日本環境変異原学会第 49 回大会, 2020/11/27, 国内（ハイブリッド開催）
13) 小島 肇:発生毒性代替法の状況と方向性について, 第 2 回日本生殖発生毒性フォーラム, 2020/11/28, 国内（web 開催）
14) Kojima H: Establishment of the Asian Consortium for Three R's,The 26th Frontier Scientists Workshop, 2020/12/18, 国外（web 開催）

15) Kojima H: Establishment of the Asian Consortium for Three R's with SAAE-I, International Webinar & 3rd National Conference of the Society for Alternatives to Animal Experiments (IWSAAE& NCSAAE-2020), 2020/12/28, 国外（web 開催）
16) Kojima H: Non-animal Alternative Toxicology and Regulatory Testing: An Update from Japan. Virtual International Conference to Mark the Launch of the Society for Alternatives to Animal Testing in Sri Lanka (SAAT-SL), 2020/2/7, 国外（web 開催）
17) Kojima H: Current projects for systemic toxicological testings in Japan. Global Collaboration Coffee When and How We Can Stop Using Animals in Toxicology, 2021 SOT Virtual Annual Meeting, 2021/3/16, 国外（web 開催）

H. 知的所有権の取得状況

1. 特許取得
   特になし
2. 実用新案登録
   特になし
3. その他
   特になし

表 1．IL-2 Luc アッセイ第三者評価委員会メンバーリスト

No.	Name	Affiliation	Country
1	Henk van Loveren	Maastricht University	Netherlands
2	Haley Neff-LaFord	Seattle Genetics, Inc.	USA
3	Barbara Kaplan	Mississippi State University	USA
4	Fujio Kayama	Jichi Medical University	Japan
5	Xingchao Geng	National Center for Safety Evaluation of Drugs (NCSED)	China
6	Sang-Hyun Kim	Kyungpook National University	Korea
7	Takao Ashikaga	National Institute of Health Sciences	Japan

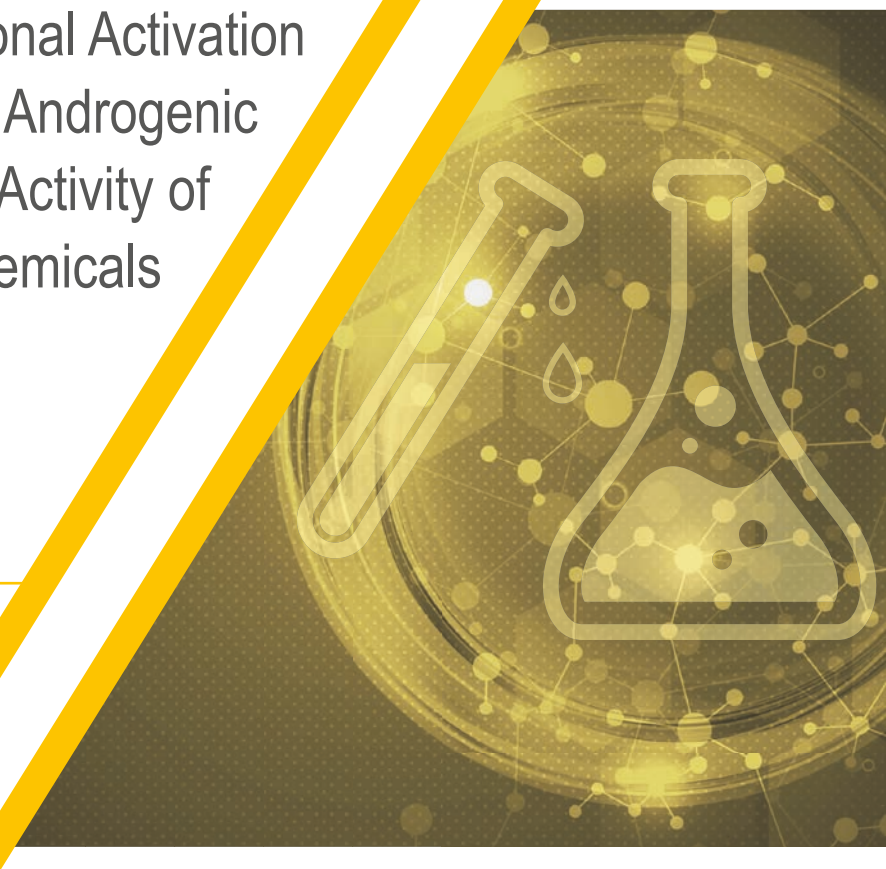
表 2．*In vitro* 免疫毒性試験　総説（Detailed Review Paper）作成メンバーリスト

No.	Name	Affiliation	Country
1	Emanuela Corsini	Università degli Studi di Milano	Italy
2	Dori Germolec	NTP/NIEHS	USA
3	Barbara Kaplan	Mississippi State University	USA
4	Henk van Loveren	Maastricht University	Netherlands
5	Haley Neff-LaFord	Seattle Genetics, Inc.	USA
6	Erwin L. Roggen	3RsMC ApS	Denmark
7	Setsuya Aiba	Tohoku University	Japan
8	Yutaka Kimura	Tohoku University	Japan
9	Takayuki Yoshimoto	Tokyo Medical University	Japan
10	Hajime Kojima	JaCVAM, National Institute of Health Sciences	Japan

Section 4
Health effects

Test Guideline No. 458

Stably Transfected Human Androgen Receptor Transcriptional Activation Assay for Detection of Androgenic Agonist and Antagonist Activity of Chemicals



26 June 2020

OECD Guidelines for the Testing of Chemicals

OECD/OCDE 458

Adopted:
26 June 2020

OECD GUIDELINE FOR THE TESTING OF CHEMICALS

Androgen Receptor TransActivation Assays for Detection of Androgenic Agonist and Antagonist Activity of Chemicals using Stably Transfected Cell Lines

Table of Contents

Introduction

Androgen receptor transactivation (ARTA) test guideline of similar in vitro methods

1. Disruption of the endocrine system may occur through a number of different mechanisms including interference with (i) hormone action mediated via nuclear receptors linked to the endocrine system (ii), hormone production via steroidogenic or other enzymes, (iii) metabolic activation or deactivation of hormones, (iv) distribution of hormones to target tissues, and (v) clearance of hormones from the body. This Test Guideline (TG) exclusively addresses transcriptional activation and inhibition of an androgen-regulated reporter gene.

2. The results of the methods in this TG should not be directly extrapolated to the complex *in vivo* situation of androgen regulation of any cellular or physiological processes.

3. This TG describes the methodology of Androgen Receptor TransActivation (ARTA) assays that detect agonist and antagonists. It comprises several mechanistically and functionally similar test methods for the identification of androgen receptor agonists and antagonists. The fully validated reference test methods described in this TG are:

- The AR-EcoScreen™ method using the AR-EcoScreen™ cell line (1) (Method 1, found in Annex C)
- The AR-CALUX® method using the AR-CALUX® cell line (2) (Method 2, found in Annex D)
- The ARTA method using the 22Rv1/MMTV_GR-KO cell line (3) (Method 3, found in Annex E)

4. These three test methods address the same endpoint, i.e. transactivation of a reporter gene by a ligand bound androgen receptor (see paragraphs 5 and 6). An overview of the similarities and differences between the test methods is given in Annex B (Tables B.1 and B.2). All three test methods are performed in 96-well plates while a high-throughput application has also been reported (but not yet validated according to OECD Guidance Document 34, 2020) for the AR-CALUX® test method (4). Method 1 includes a specificity control for the agonist detection but not for the antagonist, whereas methods 2 and 3 include a specificity control for the antagonist assay to give assurance that what is measured is a competitive antagonist. Each test method has a distinct protocol and test run acceptability criteria. Each test method has its own data interpretation criteria to conclude on agonist and antagonist activity.

Background and principles of the test methods included in this test guideline

5. *In vitro* Transactivation (TA) methods are based upon the transcription and translation of a reporter gene (e.g. luc gene) following binding of a chemical to a specific receptor and subsequent transactivation. Different reporter genes can be used in these assays. TA methods have been used to evaluate the gene expression profiles regulated by specific nuclear receptors, such as the estrogen receptors (ERs) and androgen receptors (ARs) (5) (6) (7) (8). They have been proposed for the detection of nuclear receptor-mediated transactivation (5) (6) (9).

6. Androgen agonists and antagonists act as ligands for the AR through AR binding, and may activate or inhibit the transcription of androgen responsive genes. This interaction may have the potential to trigger adverse health effects by disrupting androgen-regulated systems e.g. processes necessary for cell proliferation, normal fetal development, and reproductive function.

7. The OECD initiated a high-priority activity in 1998 to revise existing, and to develop new TGs for the screening and testing of potential endocrine disrupting chemicals. The OECD Conceptual Framework for testing and assessment of potential endocrine disrupting chemicals comprises five levels, each level corresponding to a different level of biological complexity (10). The 3 ARTA methods described in this TG are included in level 2 for *"in vitro assays providing data about selected endocrine mechanism(s)/pathway(s) (Mammalian and non mammalian methods)"*.

8. The test methods described in this TG cannot be used on their own for safety assessment decisions. They provide concentration-response data for chemicals with *in vitro* (anti)androgenic activity, which may be used for screening and prioritization purposes and can also be used as mechanistic information in a weight of evidence approach.

9. Validation studies of the AR-EcoScreen™ test method, the AR-CALUX® test method, and 22Rv1/MMTV_GR-KO test method have demonstrated their relevance and reliability (1, 3, 11).

10. An overview of the main characteristics, the acceptability criteria and the main abbreviations used in each test method is described in Annex B (Tables B.1 and B.2). For information purposes, Tables B.3a and B.3b of Annex B provide the results for the chemicals that were tested in common between at least 2 test methods of this TG. The classification comparison is made with the ICCVAM list of 2003 (6) (used as the reference list for the AR-EcoScreen™ which was adopted in 2016) and with the recently updated ICCVAM list of 2017 (12). For the antagonist testing, the 3 test method results were concordant, whereas results of agonist testing resulted in 4 non concordant classifications with the 22Rv1/MMTV_GR-KO method. A possible reason for this could be the different cell lines used in the 3 different test methods (1, 3, 11). The chemical 17β-Estradiol, known as an ER agonist, shows AR agonist activity with all 3 test methods, although in the AR-CALUX® method only a weak activity was observed.

11. Supplementary information on these chemicals as well as on an additional 13 chemicals tested with the AR-CALUX® method can be found in the validation study reports (1, 3, 11).

12. General and test method specific definitions and abbreviations used in the test methods in this TG can be found in Annex A.

Demonstration of laboratory proficiency

13. Each laboratory should demonstrate proficiency in using the test method of choice prior to using that method for testing chemicals with unknown activity. Proficiency is demonstrated by testing 8 proficiency chemicals for agonist activity (see Table B.4a in Annex B) and 9 proficiency chemicals for antagonist activity (see Tables B.4b and B.4c in Annex B). This testing will also confirm the responsiveness of the test system. Testing should be replicated at least twice, on different days, and the results should be consistent to the listed classifications and values in Tables B.4a and B.4a. Moreover, a historical database of data generated with the reference standards and the vehicle/solvent controls shall be maintained to confirm the reproducibility of the test method in the respective laboratory over time.

Test report

14. For reporting purposes, the template provided in Annex B, should be used for each test method.

Literature

1. OECD (2016), Validation report of Androgen Receptor (AR) Mediated Stably Transfected Transactivation (AR STTA) Assay to Detect Androgenic and Anti-androgenic Activities. Environment, Health and Safety Publications, Series on Testing and Assessment (No. 241), Organisation for Economic Cooperation and Development, Paris.

2. Transactivation assay for the detection of compounds with (anti)androgenic potential using AR-CALUX® cells (2019). Available at (https://tsar.jrc.ec.europa.eu/test-method/tm2010-07).

3. Validation Study Report of the 22Rv1/MMTV_GR-KO ARTA assay (2019). Available at (http://www.nifds.go.kr/brd/m_18/view.do?seq=12486&srchFr=&srchTo=&srchWord=&srchTp=&itm_seq_1=0&itm_seq_2=0&multi_itm_seq=0&company_cd=&company_nm=&page=1) .

4. Van der Burg, B., Pieterse, B., Buist, H., Lewin, G., van der Linden, S.C., Man, H.Y., Rorije, E., Piersma, A.H., Mangelsdorf, I., Wolterbeek, A.P., Kroese, E.D., van Vught-Lussemburg, B. (2015). A high troughput screening system for predicting chemically-induced reproductive organ deformities. Reprod Toxicol 55, 95-103.

5. EDSTAC (1998), Endocrine Disruptor Screening and Testing Advisory Committee (EDSTAC) Final report. Available at (https://www.epa.gov/endocrine-disruption/endocrine-disruptor-screening-and-testing-advisory-committee-edstac-final).

6. ICCVAM (2003), ICCVAM Evaluation of *In Vitro* Test Methods for Detecting Potential Endocrine Disruptors: Estrogen Receptor and Androgen Receptor Binding and Transcriptional Activation Assays. Available at (https://ntp.niehs.nih.gov/iccvam/docs/endo_docs/edfinalrpt0503/edfinrpt.pdf).

7. Jefferson, W.N., Padilla-Banks, E., Clark, G. and Newbold R. (2002). Assessing estrogenic activity of phytochemicals using transcriptional activation and immature mouse uterotrophic responses. *J. Chromat. B.*, 777, 179-189.

8. Sonneveld, E., Riteco, J.A., Jansen, H.J., Pieterse, B., Brouwer, A., Schoonen, W.G. and van der Burg, B. (2006). Comparison of *in vitro* and *in vivo* screening models for androgenic and estrogenic activities. *Toxicol. Sci.*, 89, 173-187.

9. Gray, L.E. Jr. (1998). Tiered screening and testing strategy for xenoestrogens and antiandrogens. *Toxicol. Lett.*, 102-103, 677-680.

10. OECD (2018), Revised Guidance Document No 150 on Standardised Test Guidelines for Evaluating Chemicals for Endocrine Disruption, OECD Series on Testing and Assessment, No. 150, OECD Publishing, Paris, https://doi.org/10.1787/9789264304741-en .

11. Validation Study Report on the Performance assessment of the AR-CALUX® *in vitro* method (2019). Available at (https://tsar.jrc.ec.europa.eu/test-method/tm2010-07) .

12. Kleinstreuer, N.C., Ceger, P., Watt, E.D., Martin, M., Houck, K., Browne, P., Thomas, R.S., Casey, W.M., Dix, D.J., Allen, D., Sakamuru, S., Xia, M., Huang, R., Judson, R. (2017). Development and Validation of a Computational Model for Androgen Receptor Activity. *Chem Res Toxicol*., 30(4):946-964.

Annex A. Definitions and abbreviations

General definitions and abbreviations that apply to all the test methods in this TG and/or to the tables in Annex B

Acceptability criteria: Minimum standards for the performance of experimental controls and reference standards. All acceptability criteria should be met for an experiment to be considered valid

Agonist: A chemical that binds to a specific receptor and triggers a response in the cell. It mimics the action of an endogenous ligand that binds to the same receptor

Androgen activity: The capability of a chemical to mimic a ligand in its ability to bind to and activate androgen receptors

Antagonist: A type of receptor ligand or chemical that does not provoke a biological response itself upon binding to a receptor, but blocks or dampens agonist-mediated responses

Anti-androgen activity: The capability of a chemical to suppress the action of the agonist ligand mediated through androgen receptors. AR-mediated specific anti-androgen activity can be detected in this Test Guideline.

AR: Androgen Receptor

ARE: Androgen Receptor Element

ARTA: Androgen Receptor TransActivation

BDS: BioDetection Systems (The Netherlands)

BLR: Between Laboratory Reproducibility

CERI: Chemicals Evaluation and Research Institute (Japan)

CASRN: Chemical Abstracts Service Registry Number

CRISPR-Cas9: Clustered Regularly Interspaced Short Palindromic Repeats-CRISPR associated

CV: Coefficient of Variation

Cytotoxicity: Harmful effects to cell structure or function ultimately causing cell death. It can be reflected by a reduction in the number of cells present in the well at the end of the exposure period or a reduction of the capacity for a measure of cellular function when compared to the concurrent vehicle control.

DHT: 5α-DiHydroTestosterone

DMSO: DiMethyl SulfOxide

EC_{50}: The half maximal effective concentration of a stimulating (agonist) test chemical

ED: Endocrine Disruptor

ER: Estrogen Receptor

FBS: Fetal Bovine Serum

IF: Induction Factor/Fold

InhF: Inhibition Factor/Fold

GR: Glucocorticoid Receptor

IC_{50}: The half maximal effective concentration of an inhibitory (antagonist) test chemical

InChI: International Chemical Identifier

IUPAC: International Union of Pure and Applied Chemistry

KO: KnockOut

Luc: Luciferase gene

MTA: Material Transfer Agreement

MFDS: Ministry of Food and Drug Safety (Korea)

MMTV: Mouse Mammary Tumour Virus

Negative control: Separate part of a test system treated with a chemical for which it is known that the test system should not respond. The negative control provides evidence that the test system is not responsive under the actual conditions of the assay

NIHS: National Institute of Health Sciences (Japan)

PCR: Polymerase Chain Reaction

PR: Progesterone Receptor

Positive control: Separate part of the test system treated with a chemical for which it is known that the test system should respond. The positive control provides evidence that the test system is responsive under the actual conditions of the assay

R^2: Square of the correlation coefficient (criterion for the specificity control test)

Reference chemical: A chemical used to provide a basis for comparison with the test chemical

Reference standard: Used to demonstrate the adequacy of a test method. In this TG, reference standards refer to 3 chemicals of which 2 elicit a positive response (a dose response or at one fixed concentration) and one does not provide a response. One of the 2 chemicals with a positive dose response is the reference chemical

Reliability: Measure of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating WLR and BLR

RI: Relative Induction

RLU: Relative Light Units

Run: An individual ex*periment* that evaluates chemical action on the biological outcome of the test method. Each run is a complete experiment performed on replicate wells of cells plated from a common pool of cells at the same time

SD: Standard Deviation

SMILES: Simplified Molecular-Input Line-Entry System

Study: The full range of experimental work performed to evaluate a single, specific chemical using a specific test method. In this TG, a study comprises all steps including tests of dilution of test chemical in the test media, runs (which can be pre-screen runs and comprehensive runs), data analysis, quality assurance, cytotoxicity assessments, etc. Completion of a study allows the classification of the test chemical activity on the toxicity target that is evaluated by the test method used and an estimate of potency relative to the positive reference chemical

TA: Transactivation. The initiation of mRNA synthesis in response to a specific chemical signal, such as a binding of an androgen to an androgen receptor

Test chemical: what is being tested and is not related to the applicability of the assay to the testing of mono-constituent chemicals, multi-constituent chemicals and/or mixtures

Test method: Within the context of the TG, a test method is one of the methodologies accepted as valid in meeting the performance criteria outlined in the TG. Components of the test method include, for example, the specific cell line with associated growth conditions, specific media in which the test is conducted, plate set up conditions, arrangement and dilutions of test chemicals along with any other required quality control measures and associated data evaluation steps

Test system: Any biological, chemical or physical system or a combination thereof used in a study. *In vitro* test systems are mainly biological systems (e.g. cells or tissues)

UN GHS: United Nations Globally Harmonized System of classification and labelling of chemicals

UVCBs: Chemicals of Unknown or Variable Composition, Complex Reaction Products and Biological Materials

Validated test method: A test method for which a validation study has been completed to determine the relevance (including accuracy) and reliability for a specific purpose. It is important to note that a validated test method may not have sufficient performance in terms of accuracy and reliability to be found acceptable for the proposed purpose

Validation: The process by which the reliability and relevance of a particular approach, method, process or assessment is established for a defined purpose

Vehicle control: The solvent (vehicle) that is used to dissolve test and reference standards. It is tested solely as vehicle without dissolved chemical

WLR: Within Laboratory Reproducibility

Test method specific terminology

AR-EcoScreen™ test method

AG ref: Agonist reference (500 pM of DHT) in the antagonist assay

BPA: BisPhenol A

DCC-FBS: Dextran-Coated Charcoal treated Fetal Bovine Serum

DEHP: Di(2-EthylHexyl)Phthalate

HF: HydroxyFlutamide

IC_{30}: the concentration of a test chemical at which the measured activity in an antagonist assay inhibits at level of 30% of the maximum activity induced by 500 pM DHT in each plate

PC_{AGO}: AR agonist control displaying a positive response with DHT at 10 nM

PC_{ATG}: AR antagonist control displaying a positive response with 500 pM DHT and 1 μM of HF

PC_{CT}: The response of the cytotoxic control (10 μg/mL of Cycloheximide)

PC_{10}: The concentration of a test chemical at which the response in an agonist assay is 10% of the response induced by the reference chemical (DHT at 10 nM) in each plate

PC_{50}: The concentration of a test chemical at which the response in an agonist assay is 50% of the response induced by the reference chemical (DHT at 10 nM) in each plate

PC_{max}: The concentration of a test chemical inducing the RPC_{max}

RPC_{max}: Maximum level of response induced by a test chemical, expressed as a percentage of the response induced by PC_{AGO} (10 nM DHT) on the same plate

RTA: Relative Transcriptional Activity

AR-CALUX® test method

ARE: Androgen Responsive Elements

AU: Absorbance Units

Comprehensive run: experiment carried out after the pre-screen run with a smaller dilution step (e.g. 2, 3 or 5) in order to calculate the parameters with more precision

DF: Dilution Factor

DMEM: Dulbecco's Modified Eagle's medium

EC_{10}, EC_{50}: Concentration of a test chemical at which 10% or 50% of its maximum induction response is observed

FLU: Flutamide

hAR: Human Androgen Receptor

HTS: High Throughput Screening

IATA: Integrated Approach to Testing and Assessment. IATA are pragmatic, science-based approaches for chemical hazard characterisation that rely on an integrated analysis of existing information coupled with the generation of new information using testing strategies

IC_{20}, IC_{50}: Concentration of a test chemical at which 20% or 50% inhibition is observed when compared to its maximum response

IP: Intellectual Property

LDH: Lactate Dehydrogenase

PC_{10}, PC_{50}, PC_{80}: Concentration of test chemical giving 10%, 50% or 80% induction (or inhibition) with respect to the maximum induction of the reference chemical DHT (agonist), or, the solvent control (antagonist)

PC_{max}, PC_{min}: Concentration of a test chemical where the response is maximal (corresponding to RPC_{max}) or minimal (RPC_{min})

Pre-screen run: Experiment that evaluates the dose response, usually carried out with a large dilution step (e.g. 10) in order to capture the full dose response (if possible). It serves to determine the range of concentrations to be used in a following comprehensive run.

REF RPC_{10}, REF RPC_{50}, REF RPC_{80}: Response level (as determined by relative induction (RI)) of the reference chemical DHT or Flutamide at 10%, 50% or 80%

REF EC_{50}: Concentration of the reference chemical DHT at which 50% of its maximum response is observed (in the agonist assay)

REF IC_{50}: Concentration of the reference chemical Flutamide at which 50% of its maximum response is observed (in the antagonist assay)

RPC_{max}: The maximum response level (highest induction) of the test chemical

RPC_{min}: The minimum response level (highest inhibition) of the test chemical

RI: Relative induction

Specificity control: A test which is carried out to assess if the antagonist response is the result of competitive binding to the AR

S_c: Specificity control response at a specific concentration c, expressed in relative induction.

S_c^n: Normalized specificity control response at a specific concentration c, expressed in relative induction

SC: Solvent control (agonist: assay medium plus 0.1 % solvent; antagonist: assay medium plus 0.1 % solvent and spiked with the EC_{50} concentration DHT)

VC: Vehicle control (assay medium plus 0.1% solvent, used in the antagonist assay)

Y_{ic}: Standard response at concentration c (C1-C8), expressed in relative induction) and technical replicate i (1-3)

Y_c: Average of the standard response Y_{ic} over the 3 technical replicates

ZF: Z-factor

22Rv1/MMTV_GR-KO test method

ATCC: American Type Culture Collection

Comprehensive run: experiment carried out after the pre-screen run with a smaller dilution step (e.g. 3 or 5) in order to calculate the parameters with more precision

DCC-FBS: Dextran-coated charcoal treated fetal bovine serum

DEHP: Di(2-EthylHexyl)Phthalate

GF-AFC: Glycyl phenylalanyl-aminofluorocoumarin

IC_{30}: The concentration of a chemical at which its inhibitory response equals 30% of the maximum response of the AR agonistic control (800 pM DHT) in AR antagonist assay

KTR: Korean Testing and Research Institute

NIFDS: National Institute for Food and Drug Safety Evaluation

PC_{10}: The concentration of a chemical at which its response equals 10% of the maximum response of the AR agonistic control (10 nM DHT) in AR agonist assay

PC_{50}: The concentration of a chemical at which its response equals 50% of the maximum response of the AR agonistic control (10 nM DHT) in AR agonist assay

PC_{AGO1}: Control for AR agonist assay displaying a positive response with 10 nM DHT

PC_{AGO2}: Agonist control for AR antagonist assay (800 pM DHT)

PC_{ANTA}: Antagonist control displaying a positive response with 800 pM DHT and 1 μM of Bicalutamide.

PC_{CT}: Cytotoxic control (1 mM SDS)

Pre-screen run: Experiment that evaluates the dose response, usually carried out with a large dilution step (e.g. 10) in order to capture the full dose response (if possible). It serves to determine the range of concentrations to be used in a following comprehensive run.

RTA: Relative Transcriptional Activity

SDS: Sodium Dodecyl Sulfate

Specificity control: A test which is carried out to assess if the antagonist response is the result of competitive binding to the AR

Sc: The relative induction of a test chemical at concentration c when 100 nM DHT is used in the antagonist assay (specificity control)

Yc: The relative induction of a test chemical at concentration c when 800 pM DHT is used in the antagonist assay

Annex B. Information relevant to the three test methods.

Overview tables and list of proficiency chemicals

Table B.1. Overview of the characteristics of the 3 test methods in this TG

Test method name	AR-EcoScreen™	AR-CALUX®	22Rv1/MMTV_GR-KO
Developer	Otsuka Pharmaceuticals Co., Ltd., CERI and NIHS.	BDS	MFDS, Korea Univ. and Dongguk Univ.
Cell line	AR-EcoScreen™	AR-CALUX®	22Rv1/MMTV_GR-KO
Cell type	Chinese hamster ovarian cancer cell	Human osteo-sarcoma cell	Human prostate carcinoma epithelial cell
Genetic modification	• Human AR cDNA • heat shock protein promoter -4 C3 ARE-firefly luc (*Photinus pyralis*) • SV40 promoter-renilla luc (*Renilla reniformis*) (for simultaneous measurement of cytotoxicity)	• Human AR cDNA • TATA promoter -3xARE -firefly luc (*Photinus pyralis*)	• Endogenous AR • MMTV LTR promoter containing ARE- firefly luc (*Photinus pyralis*) • Knocked out GR by CRISPR-Cas9
Special feature	• Minimal GR crosstalk due to the selection of an appropriate androgen responsive element • High throughput applicability	• No or little GR, ER and PR expression • High throughput applicability	• No ER and PR expression • GR knock-out

Availability	Material transfer agreement (MTA) including a licence agreement with Japanese Collection of Research Bioresources (JCRB) Cell Bank and cell owner	Licence agreement with BDS	Material transfer agreement (MTA) including a licence agreement with Korean Cell Bank and Korean MFDS

Table B.2a. Overview of reference standards and acceptability criteria for the three test methods for AGONIST properties.

Test method name	AR-EcoScreen™	AR-CALUX®	22Rv1/MMTV_GR-KO
AGONIST			
Reference chemical	5α-Dihydrotestosterone (DHT)	5α-Dihydrotestosterone (DHT)	5α-Dihydrotestosterone (DHT)
Criteria	Range logPC_{50} -11.03/-9.00 (log[M])	Range EC_{50} 1.10^{-10}/1.10^{-9}M	Range logPC_{50} -10.6/-9.0 (log[M])
	Range logPC_{10} -12.08/-9.87 (log[M])		Range logPC_{10} -12.2/-9.7 (log[M])
	Sigmoidal curve	Sigmoidal curve	Sigmoidal curve
	IF PC_{AGO} > 6.4 (PC_{AGO}: DHT 1.0 x 10^{-8} M)	IF DHT 1.0 x 10^{-7} M > 20	IF $PC_{AGO} \geq 13$ (PC_{AGO}: DHT 1.0 x 10^{-8} M)
	IF PC_{10} > 1 + 2SD (induction of VC)		IF PC_{10} > 1 + 2SD (induction of VC)
	CV < 20% in triplicate wells	|CV| logEC_{50} < 1.5%	
		ZF > 0.5	
Positive control	Mestanolone	17α -Methyltestosterone	Mestanolone
Criteria	Range logPC_{50} -10.15/-9.26 (log[M])	RI > 30%	Range logPC_{50} -10.2/-8.6 (log[M])
	Range logPC_{10} -10.92/-10.41 (log[M])		Range logPC_{10} -12.3/-9.8 (log[M])
	Sigmoidal curve		
	CV < 20% in triplicate wells		
Negative control	Di(2-ethylhexyl)phthalate (DEHP)	Corticosterone	Di(2-ethylhexyl)phthalate (DEHP)

Test method name	AR-EcoScreen™	AR-CALUX®	22Rv1/MMTV_GR-KO
Criteria	PC_{10} cannot be calculated	RI <10%	PC_{10} cannot be calculated
Specificity control (agonist)		NA	NA
Criteria	Confirmation by adding potent AR antagonist (1 µM HF) to clarify the non-AR mediated induction of luciferase.		

Table B.2b. Overview of reference standards and acceptability criteria for the three test methods for ANTAGONIST properties.

Test method name	AR-EcoScreen™	AR-CALUX®	22Rv1/MMTV_GR-KO
ANTAGONIST			
Reference chemical	Hydroxyflutamide (HF)	Flutamide (FLU)	Bicalutamide
Criteria	Range logIC_{50} -7.80/-6.17 (log[M])	IC_{50} range 1.0 x 10^{-7}/1.0 x 10^{-6}M	Range logIC_{50} -7.0/-5.8 (log[M])
	Range logIC_{30} -8.37/-6.41 (log[M])		Range logIC_{30} -7.5/-6.2 (log[M])
	RTA of PC_{ATG} < 46% (PC_{ATG}: 500 pM DHT + HF 1 µM))	InhF FLU 3.10^{-5}M > 10	RTA of PC_{ATG} ≤ 53.6 % (PC_{ATG}: 800 pM DHT + Bicalutamide 1 µM)
	Sigmoidal curve	Sigmoidal curve	Sigmoidal curve
	CV< 20% in triplicate wells	|CV| logIC_{50} < 3%	
		ZF > 0.5	
Positive control	Bisphenol A	Linuron	Bisphenol A
Criteria	Range log IC_{50} -7.05/-4.29 (log[M])	RI < 60%	Range log IC_{50} -6.2/-5.0 (log[M])
	Range logIC_{30} -7.52/-4.48 (log[M])		Range logIC_{30} -6.6/-5.4 (log[M])
	CV< 20% in triplicate wells		

Test method name	AR-EcoScreen™	AR-CALUX®	22Rv1/MMTV_GR-KO
Negative control	DEHP	Levonorgestrel	DEHP
Criteria	IC_{30} cannot be calculated	RI > 85%	IC_{30} cannot be calculated
Other control	IF AG_{ref} > 5 (AG_{ref}: 500 pM DHT)	NA	IF $AG_{ref} \geq 10$ (AG_{ref}: 800 pM DHT)
Specificity control (antagonist)		DHT	DHT
Criteria	NA	R^2 test chemical ≤ 0.9	R^2 test chemical < 0.9
		R^2 FLU ≤ 0.7	

NA: not applicable

Note: 1) Different mathematical techniques are used in the three methods for the calculation of IC_{50}, IC_{30} (interpolation for AR-EcoScreen™ test method and 22Rv1/MMTV_GR-KO test method; curve fitting for AR-CALUX®test method); 2) Different spiking concentrations of DHT were used in the antagonist assay: 500 pM in AR-EcoScreen™ test method; 300 pM in AR-CALUX® test method; 800 pM in 22Rv1/MMTV_GR-KO test method.

Table B.3a. Overview of results from the three test methods in this TG. Chemicals were tested in two or three methods for AGONIST properties

		Expected outcome[1]		AR-EcoScreen™			AR-CALUX®			22Rv1/MMTV_GR-KO			Chemical Class[5]	Product Class[6]
Chemical Name	CASRN	Ref. (2003)	Ref. (2017)	Outcome Validation[2]	log PC_{10}[2] (M)	log PC_{50}[2] (M)	Outcome Validation[3]	log PC_{10}[3] (M)	log EC_{50}[3] (M)	Outcome Validation[3]	log PC_{10}[4] (M)	log PC_{50}[4] (M)		
5α-Dihydrotestosterone	521-18-6	P	P	P	-12.08/-9.87	-11.03/-9.00	P	-10.64/-10.14	-9.98/-9.42	P	-10.60/-9.83	-9.73/-8.95	Steroid, nonphenolic	Pharmaceutical
Mestanolone (Methyldihydrotestosterone)	521-11-9	P		P	-10.92/-10.41	-10.15/-9.26	P	-10.26/-9.99	-9.53/-9.39	P	-10.36/-9.66	-9.65/-8.39	Steroid, nonphenolic	Pharmaceutical

Testosterone	58-22-0	P	P	P	-10.42/-9.73	-9.46/-8.96	P	-9.81/-9.60	-9.25/-8.80	P	-10.28/-9.91	-9.67/-8.66	Steroid, nonphenolic	Pharmaceutical
17β-Estradiol	50-28-2	P		P	-7.74/-6.75	-5.34/-4.88	P	-6.70/-5.85	-	P	-8.76/-8.49	-7.19/-6.03	Steroid, phenolic	Pharmaceutical
Medroxyprogesterone 17-acetate	71-58-9	P	P	P	-9.64/-8.89	-8.77/-8.37	P	-9.91/-8.32	-9.23/-7.75	P	-8.77/-8.20	-7.64/-6.01	Steroid, nonphenolic	Pharmaceutical
17α-Ethinyl estradiol	57-63-6	N		N	-		N	-		P	-6.21/-5.27	-	Steroid, phenolic	Pharmaceutical
Butylbenzyl phthalate	85-68-7	N	N	N	-		N	-		N	-		Phthalate	Plasticiser
Di(2-ethylhexyl)phthalate	117-81-7	N		N	-		N	-		N	-		Phthalate	Chemical intermediate; Plasticiser
Hydroxyflutamide	52806-53-8	N		N	-		N	-		P	-5.54/-5.04	-	Anilide	Pharmaceutical metabolite
Bisphenol A	80-05-7	N		N	-		N	-		N	-		Bisphenol	Chemical intermediate
Methyl testosterone	58-18-4	P	P		NT	NT	P	-9.73/-9.57	-9.11/-8.95	P	-10.39/-9.99	-9.63/-9.28	Steroid, nonphenolic	Pharmaceutical
Progesterone	57-83-0	P			NT	NT	N			P	-7.13/-6.19	-5.50/-5.01	Steroid, nonphenolic	Pharmaceutical
Corticosterone	50-22-6	N			NT	NT	N			P	-7.16/-5.47	,	Steroid, nonphenolic	Pharmaceutical
Levonorgestrel	797-63-7	P	P		NT	NT	P	-9.42/-9.26	-8.91/-8.61	P	-10.28/-9.73	-9.06/-8.46	Steroid, nonphenolic	Pharmaceutical
Vinclozolin	50471-44-8	N			NT	NT	N	-		N	-		Organochlorine	Pesticide
Prochloraz	67747-09-5		N		NT	NT	N	-		N	-		Imidazole	Pesticide
Atrazine	1912-24-9	N	N		NT	NT	N	-		N	-		Triazine; Aromatic amine	Pesticide
6-Propyl-2-thiouracil	51-52-5	N			NT	NT	N	-		N	-		Pyrimidines	Pharmaceutical
o,p-DDT	789-02-6	N	N		NT	NT	N	-		N	-		Organochlorine	Pesticide
Bicalutamide	90357-06-5	N			NT	NT	N	-		N	-		Anilide	Pharmaceutical

Linuron	330-55-2	P			NT	NT	N	-	N	-	Urea	Pesticide

Table B.3b. Overview of results from the three test methods in this TG. Chemicals were tested in two or three methods for ANTAGONIST properties

		Expected outcome[1]		AR-EcoScreen™			AR-CALUX®			22Rv1/MMTV_GR-KO			Chemical Class [5]	Product Class [6]
Chemical Name	CASRN	Ref. (2003)	Ref. (2017)	Outcome validation[1]	log IC_{30} [2] (M)	log IC_{50} [2] (M)	Outcome validation[3]	log PC_{80} [3] (M)	log IC_{50} [3] (M)	Outcome validation[4]	log IC_{30} [4] (M)	log IC_{50} [4] (M)		
Hydroxyflutamide	52806-53-8	P	P	P	-8.37/ -6.41	-7.80/-6.17	P	-8.63/-8.01	-7.80/-7.54	P	-8.17/-7.45	-7.79/-7.11	Anilide	Pharmaceutical metabolite
Bisphenol A	80-05-7	P	P	P	-7.52/ -4.48	-7.05/-4.29	P	-6.75/-6.12	-5.93/-5.81	P	-5.92/-5.56	-5.68/-5.29	Bisphenol	Chemical intermediate
Flutamide	13311-84-7	P		P	-6.20/ -5.69	-5.66/-5.43	P	-7.51/-6.71	-6.60/-6.23	P	-7.11/-6.62	-6.70/-6.26	Anilide	Pharmaceutical
Prochloraz	67747-09-5	P	P	P	-5.77/ -5.47	-5.44/-5.12	P	-6.42/-6.02	-5.78/-5.59	P	-6.02/-5.30	-5.47/-4.95	Imidazole	Pesticide
Vinclozolin	50471-44-8	P	P	P	-6.83/ -6.32	-6.47/-5.85	P	-7.91/-7.00	-7.50/-6.75	P	-7.22/-6.74	-6.94/-6.44	Organochlorine	Pesticide
5α-Dihydrotestosterone	521-18-6	N		N	-		N	-		N	-	-	Steroid, nonphenolic	Pharmaceutical
Mestanolone	521-11-9	N		N	-		N	-		N	-	-	Steroid, nonphenolic	Pharmaceutical
Di(2-ethylhexyl)phthalate	117-81-7	N		N	-		N	-		N	-	-	Phthalate	Chemical intermediate; Plasticiser
Atrazine	1912-24-9	N	N	N	-		N	-		N	-	-	Triazine; Aromatic amine	Pesticide
6-Propyl-2-thiouracil	51-52-5	N		N	-		N	-		N	-	-	Pyrimidines	Pharmaceutical
17β-Estradiol	50-28-2	P			NT	NT	P	-9.05/-8.04	-8.40/-7.64	P	-7.98/-7.20	-	Steroid, phenolic	Pharmaceutical

17α-Ethinyl estradiol	57-63-6	N			NT	NT	P	-8.42/-7.75	-7.57/-7.26	P	-7.91/-7.29	-7.54/-6.88	Steroid, phenolic	Pharmaceutical
Butylbenzyl phthalate	85-68-7	N			NT	NT	P	-6.13/-5.46	-5.81/-5.11	P	-5.10/-4.57	-4.86/-4.28	Phthalate	Plasticiser
Progesterone	57-83-0	P			NT	NT	P	-8.78/-8.57	-8.07/-8.03	P	-7.40/-6.30	-6.88/-5.97	Steroid, nonphenolic	Pharmaceutical
Corticosterone	50-22-6	N			NT	NT	P	-6.85/-6.77	-6.35/-6.33	P	-6.36/-6.11	-5.91-5.51	Steroid, nonphenolic	Pharmaceutical
o,p-DDT	789-02-6	P	P		NT	NT	P	-7.33/-6.84	-6.36/-6.24	P	-5.82/-5.48	-5.56/-5.21	Organochlorine	Pesticide
Bicalutamide	90357-06-5	P	P		NT	NT	P	-8.18/-7.19	-7.23/-6.69	P	-6.92/-6.37	-6.39/-6.10	Anilide	Pharmaceutical
Linuron	330-55-2	P	P		NT	NT	P	-6.64/-6.38	-5.85/-5.70	P	-5.64/-5.33	-5.33/-5.11	Urea	Pesticide
Medroxyprogesterone 17-acetate	71-58-9	N			NT	NT	N	-		N	-		Steroid, nonphenolic	Pharmaceutical
Levonorgestrel	797-63-7	N			NT	NT	N	-		N	-		Steroid, nonphenolic	Pharmaceutical

Abbreviations for Tables B.3a and B.3b: M: molar, P: Positive, N: Negative, NT: not tested

[1]Expected outcome: Classifications reported by the ICCVAM evaluation of 2003 (5) and the ICCVAM AR-reference list of 2017 (12). The 2017 reference included additional criteria for independent confirmation of reference chemical activity in at least two assays (positive) or lack of activity in at least two assays and absence of positive activity (negative) activity. Thus, some chemicals identified in the 2003 reference did not have sufficient data to meet these criteria and were excluded for the later reference.

[2]Validation report of the AR-EcoScreen™ method (minimal/maximal values of all valid runs of all participating labs).

[3]Validation report of the AR-CALUX® method (minimal/maximal values of all valid runs of all participating labs).

[4]Validation report of the 22Rv1/MMTV_GR-KO method (minimal/maximal values of all valid runs of all participating labs).

[5]Chemicals were assigned to one or more chemical classes using the U.S. National Library of Medicine's Medical Subject Headings (MeSH), an internationally recognised standardised classification scheme (available at http://www.nlm.nih.gov/mesh).

[6]Chemicals were assigned to one or more product classes using the U.S. National Library of Medicine's Hazardous Substances Data Bank (available at http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB).

Table B.3a. List of proficiency chemicals to demonstrate technical proficiency with each of the three methods in this test guideline (AGONIST assay).

			AR-EcoScreen™			AR-CALUX®			22Rv1/MMTV_GR-KO			Chemical Class [5]	Product Class [6]
Chemical Name	CASRN	AR ref. list[1]	Class[2]	log PC_{10} [2] (M)	log PC_{50} [2] (M)	Class[3]	log PC_{10} [3] (M)	log EC_{50} [3] (M)	Class[4]	log PC_{10} [4] (M)	log PC_{50} [4] (M)		
5α-Dihydrotestosterone	521-18-6	P	P	-12.08/ -9.87	-11.03/ -9.00	P	-10.64/-10.14	-9.98/-9.42	P	-10.60/-9.83	-9.73/-8.95	Steroid, nonphenolic	Pharmaceutical
Mestanolone (Methyldihydrotestosterone)	521-11-9		P	-10.92/ -10.41	-10.15/-9.26	P	-10.26/-9.99	-9.53/-9.39	P	-10.36/-9.66	-9.65/-8.39	Steroid, nonphenolic	Pharmaceutical
Testosterone	58-22-0	P	P	-10.42/ -9.73	-9.46/-8.96	P	-9.81/-9.60	-9.25/-8.80	P	-10.28/-9.91	-9.67/-8.66	Steroid, nonphenolic	Pharmaceutical
17β-Estradiol	50-28-2		P	-7.74/ -6.75	-5.34/-4.88	P	-6.70/-5.85	-	P	-8.76/-8.49	-7.19/-6.03	Steroid, phenolic	Pharmaceutical
Medroxyprogesterone 17-acetate	71-58-9	P	P	-9.64/-8.89	-8.77/-8.37	P	-9.91/-8.32	-9.23/-7.75	P	-8.77/-8.20	-7.64/-6.01	Steroid, nonphenolic	Pharmaceutical
Butylbenzyl phthalate	85-68-7	N	N	-		N	-		N	-		Phthalate	Plasticiser
Di(2-ethylhexyl)phthalate	117-81-7		N	-			-		N	-		Phthalate	Chemical intermediate; Plasticiser
Bisphenol A	80-05-7		N	-		N	-		N	-		Bisphenol	Chemical intermediate

Abbreviations: M: molar, P: Positive, N: Negative
[1]ICCVAM AR-reference list (2017) (12).

[2]Validation report of the AR-EcoScreen™ method (minimal/maximal values of all valid runs of all participating labs).
[3]Validation study report of the AR-CALUX® method (minimal and maximal values of all valid runs of all participating labs.
[4]Validation report of the 22Rv1/MMTV_GR-KO method (minimal/maximal values of all valid runs of all participating labs).
[5]Chemicals were assigned to one or more chemical classes using the U.S. National Library of Medicine's Medical Subject Headings (MeSH), an internationally recognised standardised classification scheme (available at http://www.nlm.nih.gov/mesh).
[6]Chemicals were assigned to one or more product classes using the U.S. National Library of Medicine's Hazardous Substances Data Bank (available at http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB).

Table B.4b. List of proficiency chemicals to demonstrate technical proficiency with each of the three test methods of this TG (ANTAGONIST assay).

			AR-EcoScreen™			AR-CALUX®			22Rv1/MMTV_GR-KO			Chemical Class [5]	Product Class [6]
Chemical Name	CASRN	AR ref. list[1]	Class[2]	log IC_{30} [2] (M)	log IC_{50} [2] (M)	Class[3]	log PC_{80} [3] (M)	log IC_{50} [3] (M)	Class[4]	log IC_{30} [4] (M)	log IC_{50} [4] (M)		
Hydroxyflutamide	52806-53-8		P	-8.37/-6.41	-7.80/-6.17	P	-8.63/-8.01	-7.80/-7.54	P	-8.17/-7.45	-7.79/-7.11	Anilide	Pharmaceutical metabolite
Bisphenol A	80-05-7		P	-7.52/-4.48	-7.05/-4.29	P	-6.75/-6.12	-5.93/-5.81	P	-5.92/-5.56	-5.68/-5.29	Bisphenol	Chemical intermediate
Flutamide	13311-84-7		P	-6.20/-5.69	-5.66/-5.43	P	-7.51/-6.71	-6.60/-6.23	P	-7.11/-6.62	-6.70/-6.26	Anilide	Pharmaceutical
Prochloraz	67747-09-5		P	-5.77/-5.47	-5.44/-5.12	P	-6.42/-6.02	-5.78/-5.59	P	-6.02/-5.30	-5.47/-4.95	Imidazole	Pesticide
Vinclozolin	50471-44-8		P	-6.83/-6.32	-6.47/-5.85	P	-7.91/-7.00	-7.50/-6.75	P	-7.22/-6.74	-6.94/-6.44	Organochlorine	Pesticide
Mestanolone	521-11-9		N	-		N	-		N	-		Steroid, nonphenolic	Pharmaceutical
Di(2-ethylhexyl)phthalate	117-81-7		N	-		N	-		N	-		Phthalate	Chemical intermediate; Plasticiser

Atrazine	1912-24-9		N	-	N	-	N	-	Triazine; Aromatic amine	Pesticide
6-Propyl-2-thiouracil	51-52-5		N	-	N	-	N	-	Pyrimidines	Pharmaceutical

Table B.4c. List of 2 proficiency chemicals to test the specificity of the antagonist (for the AR- CALUX® method only).

			AR-CALUX®			
Chemical Name	**CAS RN.**	**AR ref. list[1]**	**Class[3]**	**Observation in the specificity control test when testing with a lower and higher concentration of ligand DHT**	**Chemical Class [5]**	**Product Class [6]**
Ketoconazole	65277-42-1		N	Two dose responses with a shift, $R^2 \leq 0.9$	Piperazine	Pharmaceutical, Antifungal
Cycloheximide	66-81-9		N	Two dose responses with a shift , $R^2 \leq 0.9$	Piperidone	Pharmaceutical, Fungicide

Abbreviations for Tables B.4b and B.4c: M: molar, P: Positive, N: Negative

[1]ICCVAM AR-reference list (2017) (12).

[2]Validation report of the AR-EcoScreen™ method (minimal/maximal values of all valid runs of all participating labs).

[3]Validation study report of the AR-CALUX® method (minimal and maximal values of all valid runs of all participating labs.

[4]Validation report of the 22Rv1/MMTV_GR-KO method (minimal/maximal values of all valid runs of all participating labs).

[5]Chemicals were assigned to one or more chemical classes using the U.S. National Library of Medicine's Medical Subject Headings (MeSH), an internationally recognised standardised classification scheme (available at http://www.nlm.nih.gov/mesh).

[6]Chemicals were assigned to one or more product classes using the U.S. National Library of Medicine's Hazardous Substances Data Bank (available at http://toxnet.nlm.nih.gov/cgi-bin/sis/htmlgen?HSDB).

Test report

1. The following template should be used for all three test methods. The result section is test method specific.
2. The test report should include the following information:

General information:

- Name and address of the sponsor, test facility and study director;
- Reference to TG 458 and to the test method;
- Reference to the solubility method.

Demonstration of proficiency:

- Statement that the testing facility has demonstrated proficiency in the use of the test method before routine use by testing of the proficiency chemicals

Reference standards (reference chemical, positive and negative control) and test chemical:

- Source, batch/lot number, expiry date, CAS number if available;
- Purity and chemical identity of impurities as appropriate and practically feasible;
- Physical appearance, water solubility, molecular weight and additional relevant physicochemical properties to the extent available;
- Treatment prior to testing if applicable (e.g. warming);
- Storage conditions and stability to the extent available;
- Choice of solvent/vehicle for each test chemical and justification.

Mono-constituent chemical:

Chemical identification, such as IUPAC or CAS name, CAS number, SMILES or InChI code, structural formula, purity, chemical identity of impurities, as appropriate and practically feasible.

Multi-constituent chemical, UVCBs and mixtures:

Characterised as far as possible by chemical identity (see above), quantitative occurrence and relevant physicochemical properties of the constituents, to the extent available.

Solvent/vehicle:

e.g DMSO, water, ethanol;

- Source, batch/lot number;
- Justification for choice of solvent/vehicle;

Test method conditions:

- Cell line used, its source, storage and maintenance conditions, passage number and level of confluence of cells used for testing;
- Cell counting method used for seeding prior to testing and measures taken to ensure homogeneous cell number distribution;
- Luminometer used (e.g. model), including instrument settings. Luciferase substrate used (product name, supplier, lot);
- Type of plates and their supplier and code;

- Application procedure and exposure time as specified in the protocol;
- List of the acceptability criteria to be met;
- Description of any modification of the test procedure;
- Reference to the cytotoxicity procedure.

Results obtained with the AR-EcoScreen™ test method:

Tabulation of the following results for the reference standards;

- For all reference standards: normalized data of luminescent signals; results of the application of the acceptability criteria;
- Measure of error (e.g. SD, % CV or 95% confidence interval)

Tabulation of the following results for the test chemicals;

- Solubility data and stability if known;
- Measurement of precipitate in the culture medium to which the test chemical was added, as appropriate;
- For each run:
  - Cytotoxicity data;
  - Concentrations tested;
  - Normalized data of luminescent signals and a measure of error (e.g. SD, % CV or 95% confidence interval)
  - Agonist testing: PC_{max}, log PC_{10}, log PC_{50}, EC_{50} values if appropriate, the maximum fold induction level;
  - Antagonist testing: log IC_{30}, log $IC_{50,}$ the maximum fold inhibition level;
  - The result (positive or negative) per chemical after application of the decision criteria.
  - The conclusion (positive or negative) per chemical (based on the result of two or three runs)
- Number of runs performed;
- Graphs depicting the concentration-response relationship of reference chemicals and test chemicals;
- Description of any other relevant observation.

Results obtained with the AR-CALUX® test method:

Tabulation of the following results for the reference standards;

- For all reference standards: normalized data of luminescent signals; results of the application of the acceptability criteria;
- In addition, for the reference chemicals: EC_{10} and EC_{50} values for DHT, IC_{20} and IC_{50} values for Flutamide;
- Measure of error e.g. coefficient of variation

Tabulation of the following results for the test chemicals;

- Solubility data and stability if known;
- Measurement of precipitate in the culture medium to which the test chemical was added, as appropriate;
- For the pre-screen run:
  - Cytotoxicity data (LDH leakage and microscopy observations);
  - Concentrations tested;

- Normalized data of luminescent signals and a measure of error (e.g. SD, % CV or 95% confidence interval);
- For the comprehensive run:
  - Cytotoxicity data (microscopy observations);
  - Concentrations tested, including dilution factors;
  - Normalized data of luminescent signals and a measure of error (e.g. SD, % CV or 95% confidence interval);
  - Agonist testing: RPCmax, PCmax, EC10, EC50, PC10, PC50, |CV| of logEC50, if appropriate
  - Antagonist testing: RPCmin, PCmin, IC20, IC50, PC80, PC50, |CV| of logIC50, if appropriate
  - Specificity control: R2
  - The result (positive or negative) per chemical after application of the decision criteria (based on one comprehensive run).
- Graphs depicting the concentration-response relationship of reference chemicals and test chemicals; including graphs for the specificity control;
- Description of any other relevant observation.

Results obtained with the 22Rv1/MMTV_GR-KO test method:

Tabulation of the following results for the reference standards;

- For all reference standards: normalised data of luminescent signals; results of the application of the acceptability criteria;
- Measure of error (e.g. SD, % CV or 95% confidence interval)

Tabulation of the following results for the test chemicals;

- Solubility data and stability if known;
- Measurement of precipitate in the culture medium to which the test chemical was added, as appropriate;
- For each pre-screen run and each comprehensive run:
  - Cytotoxicity data;
  - Concentrations tested, including dilution factors;
  - Normalized data of luminescent signals and a measure of error e.g. coefficient of variation;
  - Agonist testing: log PC_{10}, log PC_{50}, the maximum fold induction level;
  - Antagonist testing: log IC_{30}, log $IC_{50,}$ the maximum fold inhibition level;
  - Specificity control: R^2
  - The result (positive or negative) per chemical after application of the decision criteria.
- The conclusion (positive or negative) per chemical (based on the result of two or three pre-screen runs, or, two or three comprehensive runs).
- Number of runs performed (pre-screen and comprehensive testing);
- Graphs depicting the concentration-response relationship of reference chemicals and test chemicals; including graphs for the specificity control;
- Description of any other relevant observation.

Discussion of the results

Conclusion

Annex C. (Method 1) Androgen Receptor TransActivation Assay for Detection of Androgenic Agonist and Antagonist Activity of Chemicals using the stably transfected human AR-EcoScreen™ cell line

Initial Considerations and Limitations

1. The "General Introduction" should be read before using this test method (Main body page 6-9).

2. The AR-EcoScreen™ method uses the AR-EcoScreen™ cell line to detect (anti)androgenic activity. This method exclusively addresses transactivation and inhibition of an androgen-regulated reporter gene by binding to the human AR, and therefore it should not be directly extrapolated to the complex *in vivo* situation of androgen regulation of cellular processes. In addition, the assay is only likely to inform on the activity of the parent molecule bearing in mind the limited metabolising capacities of the *in vitro* cell systems.

3. This test method is specifically designed to detect human AR-mediated transactivation and inhibition by measuring luciferase activity as the endpoint. A high-throughput assay design can be achieved by using PC values and fixed-dose format. However, chemical-dependent interference with luminescence signals are known to occur due to over-activation or inhibition of the luciferase reporter gene assay system (1) (2) (3). It is therefore possible that such interference with the luciferase reporter gene may also occur in the AR-EcoScreen™ luciferase assay systems. This should be considered when evaluating the data.

4. This cell line has been developed to have minimal glucocorticoid receptor (GR)-mediated response, however, a limitation with respect to AR selectivity is the potential for GR cross talk (4) (5). In certain cases, this may result in chemicals that activate GR being classified positive in the test method. When further investigation is deemed necessary, both non receptor-mediated luciferase signals and GR activation can be tested by incubating the test chemical with an AR antagonist (such as Hydroxyflutamide (HF)) to confirm whether the response by the test chemical is blocked or not (see Appendix 1).

5. The test method was validated using single chemicals, therefore the applicability to test mixtures has not been addressed. The test method is nevertheless theoretically applicable to the testing of multi-constituent chemicals and mixtures. Before use of the Test Guideline on a mixture for generating data for an intended regulatory purpose, it should be considered whether, and if so why, it may provide adequate results for that purpose. Such considerations are not needed, when there is a regulatory requirement for testing of the mixture.

6. Definitions and abbreviations used in this test method are described in Annex A of this TG.

7. The AR-EcoScreen™ test method has been validated by collaboration of the Chemicals Evaluation and Research Institute (CERI) and the National Institute of Health Sciences (NIHS) in Japan with support of the study management team from the OECD validation management group for non-animal testing (6).

8. The previous (2016) draft of TG 458 indicated the AR EcoScreen™ test method was run with in the antagonist assay with 0.1 µM HF. This has been updated to 1.0 µM

HF (see Table B.2b) to increase the sensitivity of the assay and align with concentrations used during the validation of the antagonist assay. Results of previous studies conducted using 0.1 µM HF are reliable if performance criteria were met.

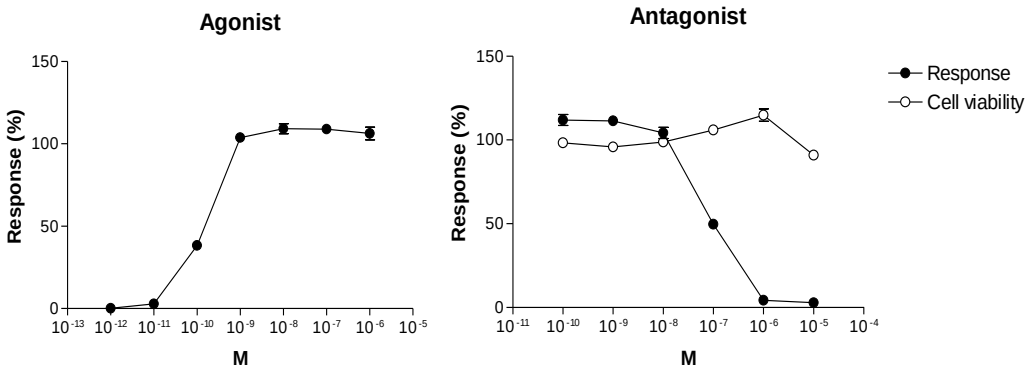
Principle of the Test Method

9. The test method using a reporter gene technique is an *in vitro* tool that provides mechanistic data. The test method is used to establish signal activation or blocking of the AR caused by a ligand. Following the ligand binding, the receptor-ligand complex translocates to the nucleus where it binds specific DNA response elements and transactivates a firefly luciferase reporter gene, resulting in an increased cellular expression of the luciferase enzyme. Luciferin is a substrate that is transformed by the luciferase enzyme to a bioluminescence product that can be quantitatively measured with a luminometer. Luciferase activity can be evaluated quickly and inexpensively with a number of commercially available test kits.

10. The test system provided in this Annex utilises the AR-EcoScreen™ cell line, which is derived from a Chinese hamster ovary cell line (CHO-K1), with three stably inserted constructs: (i) the human AR expression construct (encoding the full-length human reporter gene identical with Genbank ID of M20132 which has 21 times CAG trinucleotide short tandem repeat), and (ii) a firefly luciferase reporter construct bearing four tandem repeats of a prostate C3 gene-responsive element driven by a minimal heat shock protein promoter. The C3 gene derived androgen responsive element is selected to minimise GR-mediated responses. In addition, (iii) for cell viability assessment, a renilla luciferase reporter construct under the SV40 promoter, stably and non-inducibly expressed is transfected as to distinguish pure antagonism from a cytotoxicity-related decrease of luciferase activity. The two enzyme activities can be measured simultaneously in the same cell and in the same well. This feature facilitates the detection of the antagonist (7) (8). AR-Ecoscreen GR KO M1(JCRB1761), a modified cell line with the GR KO via genome editing and not yet officially validated, is available from JCRB for more stringent investigation of AR-mediated antagonism.(9).

11. Data interpretation for an **AR agonistic effect** is based upon the maximum response level induced by a test chemical. If this response equals or exceeds 10% of the response induced by 10 nM 5α-Dihydrotestosterone (DHT), the AR agonist control (PC_{AGO}) (i.e. the log PC_{10}), the test chemical is considered positive. Data interpretation for an **AR antagonistic effect** of a test chemical is based on a cut-off of a 30% inhibitory response against 500 pM DHT (i.e. the log IC_{30}). If the response exceeds this 30% AR blocking, then the chemical is considered a positive AR antagonist. Data analysis and interpretation are discussed in greater detail in paragraphs 48-60. Typical representations of the agonist and antagonist reference chemical curves (DHT and HF) are shown in Figure C.1.

Figure C.1. Typical positive control responses



Demonstration of Laboratory Proficiency

12. Prior to testing chemicals with unknown activity in the AR EcoScreen™, the responsiveness of the test system should be confirmed by each laboratory to yield the expected results, at least once for each newly prepared batch of cell stocks taken from the frozen stock. This is done by independently testing the proficiency chemicals listed in the Tables B.4a for agonism and B.4b for antagonism of Annex B in this TG. This should be done at least in duplicate, on different days, and the results should be consistent to the classifications and values of the Tables B.4a and B.4b in Annex B and any deviations should be justified. However, the proficiency substances are classified in Tables B.2a and B.2b by their known predominant activity which should be used for proficiency evaluation.

Procedure

Cell lines

13. The stably transfected AR-EcoScreen™ cell line should be used for the assay. The cell line can be obtained from the Japanese Collection of Research Bioresources (JCRB) Cell Bank as reference No. JCRB1328, upon signing a Material Transfer Agreement (MTA) including license agreement.

14. Only cells characterised as mycoplasma-free should be used in testing. PCR based methods are recommended for a sensitive detection of mycoplasma infection (10) (11) (12).

Stability of the cell line

15. To monitor the stability of the cell line for the **agonist assay**, DHT, Mestanolone and Di(2-ethylhexyl)phthalate (DEHP) should be used as reference standards. A complete concentration response curve for all three reference standards, at the test concentration range provided in Table C.1b and the plate concentration assignment shown in Table C.a, should be obtained at least once each time the assay is performed, and the results should be in agreement with the results provided in Tables C.1a and C.1b.

16. To monitor the stability of the cell line for measuring **AR antagonism**, HF, Bisphenol A (BPA) and DEHP should be used as reference standards. A complete concentration response curve for all three reference standards, at the test concentration range provided in Table C.1d and the plate concentration assignment shown in Table C.2b,

should be obtained at least once each time the assay is performed, and the results should be in agreement with the results provided in Tables C.1c and C.1d.

Cell culture and plating conditions

17. The following mediums should be prepared:

- Medium for dilution: Phenol Red Free D-MEM/F-12.
- Medium for cell propagation: Phenol Red Free D-MEM/F-12 supplemented with 5% v/v fetal bovine serum (FBS), Zeocin (200 μg/mL), Hygromycin (100 μg/mL), Penicillin (100 units /mL), and Streptomycin (100 μg/mL).
- Medium for the assay plate: Phenol Red Free D-MEM/F-12 supplemented with 5%v/v Dextran-coated charcoal treated (DCC)-FBS, Penicillin (100 units/mL), and Streptomycin (100 μg/mL).

18. Cells should be maintained in a CO_2 incubator (5% CO_2) at 37±1˚C with medium for cell propagation. Upon reaching 75-90% confluency (i.e. every 3-4 days), cells are subcultured to 10 mL at a density of 0.4-0.8 × 10^5 cell/mL in φ100 mm cell culture dishes. To prepare the assay plate (96-well plate), cells should be suspended in the medium for the assay plate and then plated into wells of a microtiter plate containing 90 μL/well at a density of 1.0 × 10^5 cells/mL. Next, the cells should be pre-incubated in a 5% CO_2 incubator at 37˚±1˚C for 24 hours before chemical exposure.

19. To maintain the integrity of the response, the cells should be grown for more than one passage from the frozen stock in the conditioned media for cell propagation and should not be cultured for more than 40 passages. The AR-EcoScreen™ cell line will be stable up to three months under suitable culture condition.

20. The DCC-FBS can be obtained from commercial sources. The selection of DCC-FBS is critical for the assay performance; therefore, the appropriate DCC-FBS should be selected based on the proliferative capacity and confirmation of effect on assay performance with the reference standards.

Acceptability criteria

Positive and negative reference standards

21. Prior to, and during the study, the responsiveness of the test system should be verified using the appropriate concentrations of known reference standards provided in Tables C.1b and C.1d, with DHT and Mestanolone as the positive reference standards for the agonist assay, HF and BPA as the positive reference standards for the antagonist assay, and DEHP as the negative reference standard for the agonist and antagonist assay. Acceptable range values derived from the validation study are also given in Table C.1b and Table C.1c (2). These three concurrent reference standards for each AR agonist/antagonist assay should be included in every AR agonist/antagonist experiment (conducted under the same conditions including the materials, passage level of cells and by the same technicians), and the results should fall within the given acceptable limits and the shape of concentration-response curve of positive reference standards should be sigmoidal. If this is not the case, the cause for the failure to meet the acceptability criteria should be determined (e.g. cell handling, quality of serum and antibiotics, concentration, etc.) and the assay repeated. Once the acceptability criteria have been achieved, it is essential in order to ensure

minimum variability of log PC_{50}, log PC_{10}, log IC_{30}, log IC_{50} values, that the use of materials for cell culturing is consistent.

22. The acceptability criteria of three concurrent reference standards can ensure the accuracy of quantitative sensitivity of the assay, but for the purposes of qualitative assessment, deviations from acceptable ranges of the reference standards (as specified in Tables C.1b and C.1c) could be allowed if the quality criteria are met. However the reference standards should be included with each experiment and the results should be judged according to the parameters indicated in Tables C.1b and C.1c and the concentration-response curve of the positive reference standards (reference chemical and positive control) should be sigmoidal.

Table C.1a. Quality criteria for AR agonist assay

Induction fold of PC_{AGO} (10 nM DHT)	$\geq$ 6.4
IF PC_{10}	Greater than 1 +2SD (induction of VC)

IF PC_{10}: induction fold corresponding to the PC_{10} (10%) of AR agonist control (PC_{AGO}:10 nM of DHT)
SD: Standard Deviation, VC: Vehicle Control

Induction fold of PC_{AGO} is calculated by the following equation:

$$\text{Induction fold of } PC_{AGO} = \frac{\text{Mean RLU of } PC_{AGO} \text{ (10 nM DHT)}}{\text{Mean RLU of VC}}$$

RLU: Relative Light Units

Table C.1b. Acceptable range of the reference standards for AR agonist assay

Chemical Name [CAS RN]	Judgment	logPC_{10}	logPC_{50}	Test range
5α-Dihydrotestosterone (DHT)[521-18-6]	Positive: PC_{10} should be calculated	-12.08 ~- 9.87	-11.03 ~ -9.00	1.0×10^{-12} ~ 1.0×10^{-6}M
Mestanolone[521-11-9]	Positive: PC_{10} should be calculated	-10.92 ~- 10.41	-10.15 ~ -9.26	1.0×10^{-12} ~ 1.0×10^{-6}M
Di(2-ethylhexyl)phthalate (DEHP) [117-81-7]	Negative: PC_{10} should not be calculated	-	-	1.0×10^{-11} ~ 1.0×10^{-5}M

Table C.1c. Quality criteria for AR antagonist assay

Induction fold of AG ref	≥ 5.0
RTA of PC_{ATG} (%)	≤46

AG ref = Agonist reference (500 pM DHT) in the antagonist assay
RTA : Relative Transcriptional Activity
PC_{ATG} = AR Antagonist control (500pM DHT, 1 μM HF)

Induction fold of AG ref is calculated by the following equation:

$$\text{Induction fold of AG ref} = \frac{\text{Mean RLU of AG ref (500 pM DHT)}}{\text{Mean RLU of VC}}$$

VC: Vehicle Control, RLU: Relative Light Units

RTA of PC_{ATG} (%) is calculated by the following equation;

$$\text{RTA of } PC_{ATG}\,(\%) = \left(\frac{\text{RLU of } PC_{ATG}\text{-Mean RLU of VC}}{\text{Mean RLU of AG ref - Mean RLU of VC}}\right) \times 100$$

Table C.1d. Acceptable range of the reference standards for AR antagonist assay

Chemical Name [CAS RN]	Judgment	log IC_{30}	log IC_{50}	Test range
Hydroxyflutamide (HF) [52806-53-8]	Positive: IC30 should be calculated	-8.37 ~ -6.41	-7.80 ~ -6.17	1.0×10^{-10} ~ 1.0×10^{-5}M
Bisphenol A (BPA) [80-05-7]	Positive: IC30 should be calculated	-7.52 ~ -4.48	-7.05 ~ -4.29	1.0×10^{-10} ~ 1.0×10^{-5}M
Di(2-ethylhexyl)phthalate (DEHP) [117-81-7]	Negative: IC30 should not be calculated	-	-	1.0×10^{-10} ~ 1.0×10^{-5}M

Vehicle control, AR agonist control and AR antagonist control

23. **For the agonist assay**, AR agonist control (PC_{AGO}) wells (n=4) treated with an endogenous ligand (10 nM of DHT), vehicle control (VC) wells (n=4) treated with vehicle alone and positive control for cytotoxicity (PC_{CT}, 10 μg/mL of Cycloheximide) wells (n=4) should be prepared on each assay plate in accordance with the plate design indicated in Table C.2a and Table C.3b.

24. **For the antagonist assay**, vehicle control (n=3), AR agonist control (PC_{AGO}, 10 nM of DHT, n=3), AR antagonist control control (PC_{ATG}, 500 pM DHT and 1 μM of HF, n=3), positive control for cytotoxicity (PC_{CT}, 10 μg/mL of Cycloheximide, n=3) and agonist reference (AG ref, 500 pM of DHT, n=12) should be set-up at each assay plate in accordance with the plate design indicated in Table C.2b and Table C.3b.

Quality criteria for AR agonist assay

25. The mean luciferase activity of the PC_{AGO} (10 nM DHT) should be equal to or higher than 6.4-fold compared with the mean VC on each plate for the agonist assay. These criteria were established based on the reliability of the endpoint values from the validation study.

26. With respect to the quality control of the assay, the induction fold corresponding to the log PC_{10} (10%) of the AR agonist control (PC_{AGO}: 10 nM of DHT) (IF PC_{10}) should be greater than 1+2SD of the induction value (=1) of the concurrent VC.

Quality criteria for AR antagonist assay

27. The mean luciferase activity of the AG ref (500 pM DHT) should be equal to or higher than 5.0-fold compared with the mean VC on each plate for antagonism assay. These criteria were established based on the reliability of the endpoint values from the validation study.

28. RTA of PC_{ATG} (500 pM DHT and 1 μM HF) should be less than 46%.

In summary:

29. Acceptability criteria are the following:

For AR agonist assay:

- The mean luciferase activity of the PC_{AGO} (10 nM DHT) should be equal to or higher than6.4-fold compared with the mean VC on each plate.
- The induction fold corresponding to the log PC_{10} value of the concurrent PC_{AGO} (10 nM DHT) should be greater than 1+2SD of the induction fold value of the VC.
- The shape of concentration-response curve of positive reference standards should be sigmoidal.
- The results of the three reference standards should be within the acceptable range (Table C.1b).

For AR antagonist assay:

- Induction fold of AG ref ([500 pM DHT]/[Vehicle Control]) should be equal to or higher than 5.0 compared with the mean VC on each plate.
- RTA of PC_{ATG} (%) should be less than 46.
- The shape of concentration-response curve of positive reference standards should be sigmoidal.
- The results of the three reference standards should be within the acceptable range (Table C.1d).

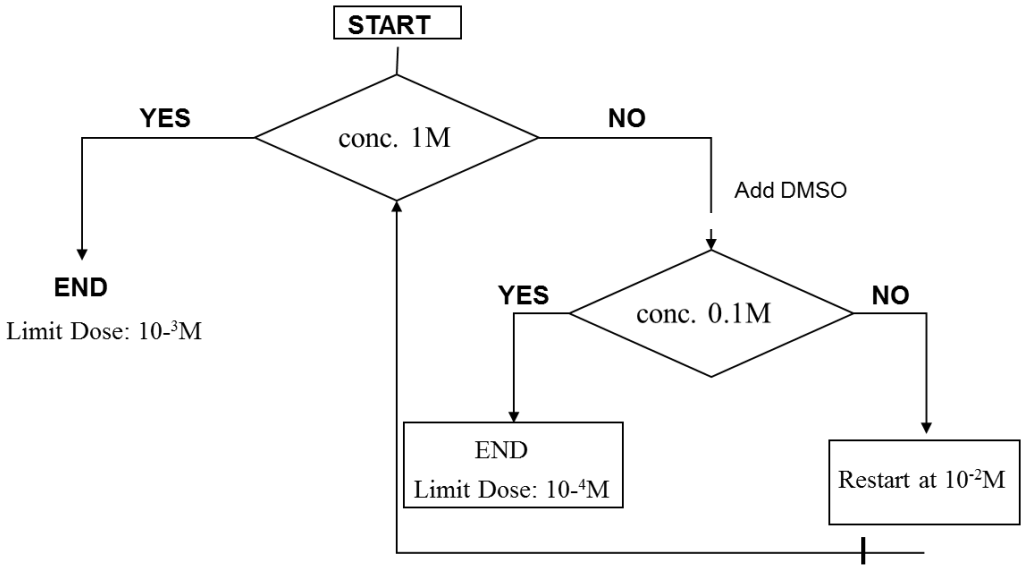
Vehicle

30. An appropriate solvent should be used as the concurrent VC at the same concentration for the different positive and negative controls and the test chemicals. Test chemicals should be dissolved in a solvent that solubilises the test chemical and is miscible with the cell medium. Water, ethanol (95% to 100% purity) and dimethyl sulfoxide (DMSO) may be suitable vehicles accepted by the cells. Generally, DMSO is used. In this case, the final concentration in the well should not exceed 0.1% (v/v). For any other vehicle (e.g. ethanol), it should be demonstrated that the maximum concentration used is not cytotoxic and does not interfere with the assay performance (as confirmed by response of renilla luciferase).

Preparation of test chemicals

31. The test chemicals should be dissolved in an appropriate solvent (see paragraph 30) and serially diluted with the same solvent at a common ratio of 1:10. In order to define the highest soluble concentration of the test chemical, a solubility test should be carried out following the flow diagram shown in Figure C.2.

Figure C.2. Diagram for solubility test



Limit dose: the highest concentration to be tested as the assay concentration.
YES: No precipitation, NO: Precipitation

32. A solubility test is a very important step to determine the maximum concentration for the assay and it may affect the sensitivity of the assay. Maximum concentration should be selected based on the avoidance of precipitation at highest concentration ranges in a medium for the assay plate. Precipitation observed at any concentration should be noted, but these concentrations should not be included in the dose-response analysis.

Cytotoxicity evaluation

33. For AR antagonists, the presence of increasing levels of cytotoxicity can significantly alter or eliminate the typical sigmoidal response and should be considered when interpreting the data. Cytotoxicity can be evaluated with renilla luciferase activity in the AR-EcoScreen™ cell line, which was originally established to express renilla luciferase constitutively. Accordingly, AR-mediated transcriptional activity and cytotoxicity should be evaluated simultaneously in the same assay plate. For AR agonists, cytotoxicity can also affect the shape of a concentration response curve. In such case, evaluation of cytotoxicity should be performed or evaluated from the results of antagonist assay conducted for same test chemical.

34. Should the results of the cytotoxicity test show that the concentration of the test chemical has reduced renilla luciferase activity by 20% or more, this concentration is regarded as cytotoxic, and the concentrations at or above the cytotoxic concentration should be excluded from the evaluation. The maximum concentration should be considered to be reduced when intrinsic cytotoxic effect is observed at the result of initial run of the test chemical. Cytotoxicity (%) of each well is calculated by the following equations and the mean of triplicate wells of same concentration is calculated for the cytotoxicity (%) of each concentration of test chemicals.

For the agonist assay;

$$\text{Cytotoxicity}\,(\%) = 100 - \left(\frac{\text{RLU of each well-Mean RLU of } PC_{CT}}{\text{Mean RLU of VC - Mean RLU of } PC_{CT}} \right) \times 100$$

For the antagonist assay;

$$\text{Cytotoxicity}\,(\%) = 100 - \left(\frac{\text{RLU of each well-Mean RLU of } PC_{CT}}{\text{Mean RLU of AG ref - Mean RLU of } PC_{CT}} \right) \times 100$$

Test chemical exposure and assay plate organisation

35. For the AR agonist assay, each test chemical should be serially diluted in DMSO or appropriate solvent, by using a single column of polypropylene plate or other appropriate item, and added to the wells of a microtiter plate to achieve final serial concentrations in the assay, from the maximum concentration determined by the solubility test with common dilution ratio of 10 (for example 1 mM, 100 μM, 10 μM, 1 μM, 100 nM, 10 nM and 1 nM [10^{-3}-10^{-9} M]) for triplicate testing.

36. For each test concentration of the test chemical, the procedure for chemical dilutions (Steps 1 and 2) and for exposing the cells (Step 3) can be conducted as follows:

- Step 1: Chemical dilution: First dilute 10 μL of the test chemical in solvent into 90 μL of media.
- Step 2: Then 10 μL of the diluted chemical prepared in Step 1 should be diluted into 90 μL of the media.
- Step 3: Chemical exposure of the cells: Add 10 μL of diluted chemical solution (prepared in Step 2) to an assay well containing 9×10^3 cells/90 μL/well.
    - The recommended final volume of media required for each well is 100 μL.

37. Reference standards and test samples can be assigned as shown in Table C.2a and Table C.3a.

Table C.2a Example of plate concentration assignment of the reference standards in the assay plate for the agonist assay

Row	DHT			Mestanolone			DEHP			Test Chemical[#]		
	1	2	3	4	5	6	7	8	9	10	11	12
A	1 µM	→	→	1 µM	→	→	10 µM	→	→	1 mM	→	→
B	100 nM	→	→	100 nM	→	→	1 µM	→	→	100 µM	→	→
C	10 nM	→	→	10 nM	→	→	100 nM	→	→	10 µM	→	→
D	1 nM	→	→	1 nM	→	→	10 nM	→	→	1 µM	→	→
E	100 pM	→	→	100 pM	→	→	1 nM	→	→	100 nM	→	→
F	10 pM	→	→	10 pM	→	→	100 pM	→	→	10 nM	→	→
G	1 pM	→	→	1 pM	→	→	10 pM	→	→	1 nM	→	→
H	VC	→	→	→	PC_{AGO}	→	→	→	PC_{CT}	→	→	→

VC: Vehicle control (DMSO);
PC_{AGO}: AR agonist control (10 nM of DHT);
PC_{CT}: Cytotoxicity control (10 µg/mL of Cycloheximide);
#: concentration of test chemical is an example

38. For the AR antagonist assay, each test chemical should be serially diluted in DMSO or appropriate solvent by using a single column of polypropylene plate or other appropriate item, and added to the wells of a microtiter plate to achieve final serial concentrations in the assay, from the maximum concentration determined by the solubility test with common dilution ratio of 10 (for example 1 mM, 100 µM, 10 µM, 1 µM, 100 nM, and 10 nM [1.0×10^{-3}-1.0×10^{-8} M]) for triplicate testing.

39. For each test concentration of the test chemical the procedure for chemical dilutions (Steps 1 and 2) and for exposing cells (Step 3) can be conducted as follows:

- Step 1: Chemical dilution: First dilute 10 µL of the test chemical in the solvent to a volume of 90 µL media containing 56 nM DHT/DMSO*.
- Step 2: Then 10 µL of the diluted chemical prepared in Step 1 should be diluted into 90 µL of the media.
- Step 3: Chemical exposure of the cells: Add 10 µL of diluted chemical solution (prepared in Step 2) to an assay well containing 9×10^3 cells/90 µL/well.
    - The recommended final volume of media required for each well is 100 µL.

* 56 nM DHT/DMSO is added to achieve 500 pM DHT, 0.1% DMSO after dilution.

40. Reference standards and test samples can be assigned as shown in Table C.2b and Table C.3b.

Table C.2b. Example of plate concentration assignment of the reference standards in the assay plate for the antagonist assay

Row	HF			Bisphenol A			DEHP			Test chemical[#]		
	1	2	3	4	5	6	7	8	9	10	11	12
A	10 µM	→	→	10 µM	→	→	10 µM	→	→	1 mM	→	→
B	1 µM	→	→	1 µM	→	→	1 µM	→	→	100 µM	→	→
C	100 nM	→	→	100 nM	→	→	100 nM	→	→	10 µM	→	→
D	10 nM	→	→	10 nM	→	→	10 nM	→	→	1 µM	→	→
E	1 nM	→	→	1 nM	→	→	1 nM	→	→	100 nM	→	→
F	100 pM	→	→	100 pM	→	→	100 pM	→	→	10 nM	→	→
G	AG ref	→	→	→	→	→	→	→	→	→	→	→
H	VC	→	→	PC_{AGO}	→	→	PC_{ATG}	→	→	PC_{CT}	→	→

VC: Vehicle control (DMSO);
PC_{AGO}: AR agonist control (10 nM of DHT);
AG ref: AR agonist reference (DMSO)
PC_{ATG}: AR antagonist control (1 µM of HF);
PC_{CT}: Cytotoxicity control (10 µg/mL of Cycloheximide);
** Gray colored wells are spiked with 500pM DHT
#: concentration of test chemical is an example

41. The reference standards (DHT, Mestanolone and DEHP for the agonist assay; HF, BPA and DEHP for the antagonist assay) should be tested in every experiment (as indicated in Tables C.2a and C.2b). Wells treated with 10 nM of DHT (PC_{AGO}), wells treated with DMSO (or appropriate solvent) alone (VC) should be included in each test assay plate for the agonist assay as well as a cytotoxicity control with 10 µg/mL of Cycloheximide (PC_{CT}) (Table C.3a). In the case of the antagonist assay, a AR agonist control with 10 nM of DHT (PC_{AGO}), an AR agonist reference with DMSO (or appropriate solvent) and spiked 500 pM DHT(AG ref), a AR antagonist control with 1 µM of HF and spiked 500 pM DHT (PC_{ATG}) and cytotoxicity control with 10 µg/mL of Cycloheximide (PC_{CT}) should be included in each assay plate (Table C.3b). If cells from different sources (e.g. different passage number, different lot numbers, etc.,) are used in the same experiment, the reference standards should be tested for each cell source.

Table C.3a. Example of plate concentration assignment of test chemicals and plate control chemicals in the assay plate for agonist assay

Row	Test Chemical 1			Test Chemical 2			Test Chemical 3			Test Chemical 4		
	1	2	3	4	5	6	7	8	9	10	11	12
A	conc 1 (10 μM)	→	→	1 mM	→	→	1 μM	→	→	10 nM	→	→
B	conc 2 (1 μM)	→	→	100 μM	→	→	100 nM	→	→	1 nM	→	→
C	conc 3 (100 nM)	→	→	10 μM	→	→	10 nM	→	→	100 pM	→	→
D	conc 4 (10 nM)	→	→	1 μM	→	→	1 nM	→	→	10 pM	→	→
E	conc 5 (1 nM)	→	→	100 nM	→	→	100 pM	→	→	1 pM	→	→
F	conc 6 (100 pM)	→	→	10 nM	→	→	10 pM	→	→	0.1 pM	→	→
G	conc 7 (10 pM)	→	→	1 nM	→	→	1 pM	→	→	0.01 pM	→	→
H	VC	→	→	→	PC_{AGO}	→	→	→	PC_{CT}	→	→	→

VC: Vehicle control (DMSO);
PC_{AGO}: AR agonist control (10 nM of DHT);
PC_{CT}: Cytotoxicity control (10 μg/mL of Cycloheximide);
The concentration of test chemicals is provided as an example.

Table C.3b. Example of plate concentration assignment of test chemicals and plate control chemicals in the assay plate for antagonist assay

Row	Test Chemical 1			Test Chemical 2			Test Chemical 3			Test Chemical 4		
	1	2	3	4	5	6	7	8	9	10	11	12
A	conc 1 (10 µM)	→	→	1 mM	→	→	1 µM	→	→	10 nM	→	→
B	conc 2 (1 µM)	→	→	100 µM	→	→	100 nM	→	→	1 nM	→	→
C	conc 3 (100 nM)	→	→	10 µM	→	→	10 nM	→	→	100 pM	→	→
D	conc 4 (10 nM)	→	→	1 µM	→	→	1 nM	→	→	10 pM	→	→
E	conc 5 (1 nM)	→	→	100 nM	→	→	100 pM	→	→	1 pM	→	→
F	conc 6 (100 pM)	→	→	10 nM	→	→	10 pM	→	→	100 pM	→	→
G	AG ref	→	→	→	→	→	→	→	→	→	→	→
H	VC	→	→	PC_{AGO}	→	→	PC_{ATG}	→	→	PC_{CT}	→	→

VC: Vehicle control (DMSO);
PC_{AGO}: AR agonist control (10 nM of DHT);
AG ref: AR agonist reference (DMSO)
PC_{ATG}-: AR antagonist control (1 µM of HF);
PC_{CT}: Cytotoxicity control (10 µg/mL of Cycloheximide);
** Gray colored wells are spiked with 500pM DHT
The concentration of test chemicals is provided as an example.

43. The lack of edge effects should be confirmed, as appropriate, and if edge effects are suspected, the plate layout should be altered to avoid such effects. For example, a plate layout excluding the edge wells can be employed.

44. After adding the chemicals, the assay plates should be incubated in a 5% CO_2 incubator at 37±1ºC for 20-24 hours to induce the reporter gene products.

45. Special considerations will need to be applied to those chemicals that are highly volatile. In such cases, nearby control wells may generate false positives, and this should be considered in light of expected and historical control values. In the few cases where volatility may be of concern, the use of "plate sealers" may help to effectively isolate individual wells during testing, and is therefore recommended in such cases.

46. Repetition of definitive tests for the same chemical should be conducted on different days using freshly prepared assay reagents and dilutions of the test chemicals, to ensure independence. In cases where multiple chemicals are concurrently tested within a

single run, maintaining the same plate design, while changing the order in which chemicals are added to the test wells, would be preferable to avoid the effects of location of chemical.

Luciferase activity measurements

47. A commercial dual-reporter assay system (e.g. Promega, E2920 or its equivalents) is preferable to detect both of the AR response (firefly luciferase activity) and cytotoxicity (renilla luciferase activity) simultaneously, as long as the acceptability criteria are met. The assay reagents should be selected based on the sensitivity of the luminometer to be used. Procedure should be followed according to the manufacturer's instructions with the following modifications For instance, when using Dual-Glo Luciferase Assay system (Promega, E2920), 60 µL of supernatant should be removed from a well of assay plate before adding the substrate, then 40 µL of the first substrate should be directly added into the assay wells and measure the firefly luciferase signal. And finally add 40 µL of the second substrate into the assay wells of the original plate to detect renilla luciferase activity. A luciferase assay reagent [e.g. Steady-Glo® Luciferase Assay System (Promega, E2510, or equivalents)] or a standard luciferase assay system (Promega, E1500, or equivalents) can be used to detect only for the AR response (firefly luciferase activity). When using Steady-Glo Luciferase Assay System (Promega, E2510), 40 µL of prepared reagent should be directly added into the assay wells. When using a standard luciferase assay system (Promega, E1500, or equivalents), the substrate should be added after adding the Cell Culture Lysis Reagent (Promega, E1531, or equivalents).

Analysis of Data

48. **For the Agonist assay**, to obtain the relative transcriptional activity to the positive control (10 nM DHT), the luminescence signals from the same plate can be analysed according to the following steps (other equivalent mathematical processes are also acceptable):

- Step 1. Calculate the mean value for the vehicle control (VC).
- Step 2. Subtract the mean value of the VC from each well value in order to subtract any vehicle-driven effect or noise.
- Step 3. Calculate the mean for the corrected PC_{AGO} (= the normalised PC_{AGO}).
- Step 4. Divide the corrected value of each well in the plate by the mean value of the normalised PC_{AGO} (PC_{AGO} is set to 100%).
- The final value of each well is the relative transcriptional activity for that well compared to the PC_{AGO} response.
- Step 5. Calculate the mean value of the relative transcriptional activity for each concentration of the test chemical. There are two dimensions to the response: the averaged transcriptional activity (response) and the concentration at which the response occurs (see paragraphs 51-60).

49. **For the Antagonist assay**, to obtain the relative transcriptional activity, the luminescence signals from the same plate can be analysed according to the following steps (other equivalent mathematical processes are also acceptable):

- Step 1. Calculate the mean value for the VC.
- Step 2. Subtract the mean value of the VC from each well value in order to subtract any vehicle-driven effect or noise.
- Step 3. Calculate the mean for the corrected AG ref (= the normalised AG ref).

- Step 4. Divide the corrected value of each well in the plate by the mean value of the normalised the AG ref (AG ref is set to 100%).

50. The final value of each well is the relative transcriptional activity for that well compared to the maximum response of the AG ref.

- Step 5. Calculate the mean value of the relative transcriptional activity for each concentration group of the test chemical. There are two dimensions to the response: the averaged transcriptional activity (response) and the concentration at which the response occurs (see paragraphs 51-60).

Calculation of parameters: EC_{50}, log PC_{50}, log PC_{10}, log IC_{50} and log IC_{30} induction considerations

51. The full concentration-response curve is required for the calculation of the EC_{50}, but this may not always be achievable or practical due to limitations of the test concentration range (for example due to cytotoxicity or solubility problems). However, as the EC_{50} and maximum induction level (corresponding to the top value of the Hill-equation) are informative parameters, these parameters should be reported where possible. For the calculation of EC_{50} and maximum induction level, appropriate statistical software should be used (e.g. Graphpad Prism statistical software).

52. If the Hill's logistic equation is applicable to the concentration response data, the EC_{50} should be calculated by the following equation (13):

- $Y=Bottom + (Top-Bottom) / (1+10 \exp ((\log EC_{50} -X) \times Hill\ slope))$
- Where: X is the logarithm of concentration;
- and, Y is the response and Y starts at the Bottom and goes to the Top in a sigmoid curve. Bottom is fixed at zero in the Hill's logistic equation.

53. To evaluate cytotoxicity, cell viability should be expressed as the percentage of renilla luciferase activity of the chemically-treated wells to the mean renilla luciferase activity of the wells of the vehicle control for the agonist assay or the mean renilla luciferase activity of the wells of AG ref (500 pM DHT) for the antagonist assay, in accordance with equations indicated in paragraph 34.

54. In the case of the agonist assay, the following information should be provided for each test chemical:

(i) The maximum level of response induced by a test chemical, expressed as a percentage against the response induced by PC_{AGO}(10 nM DHT) on the same plate (RPC_{max}).

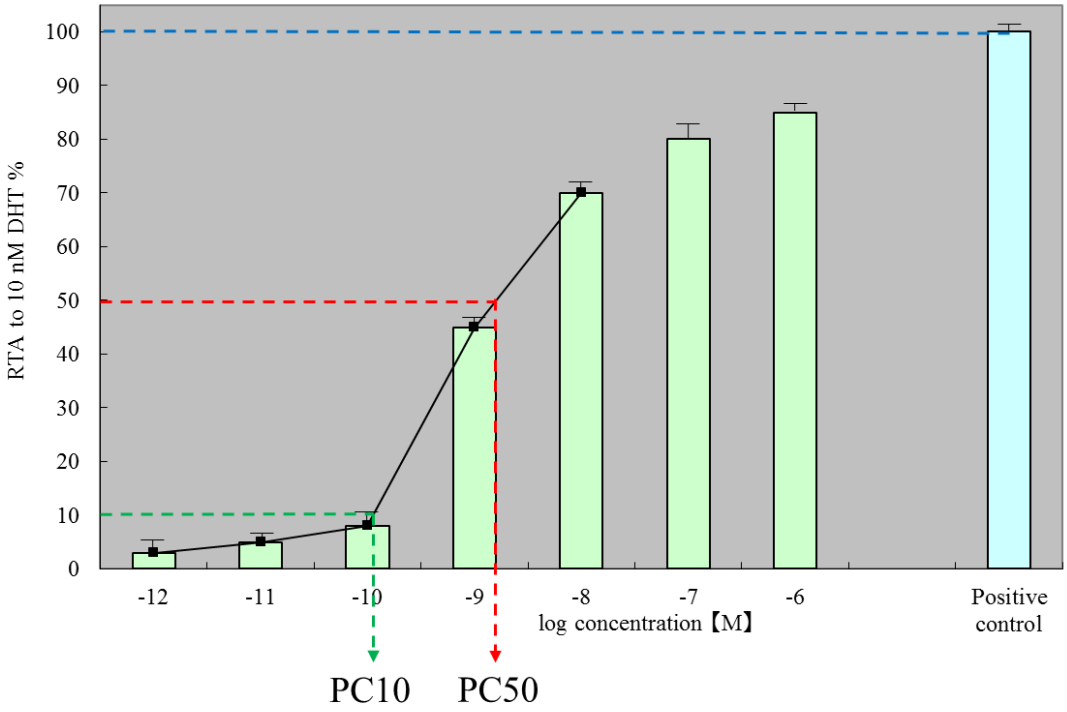
(ii) For positive chemicals, the concentrations that induce an effect corresponding to that of a 10% effect for the reference chemical DHT (log PC_{10}) and, if appropriate, to 50% effect for the reference chemical DHT (log PC_{50}).

55. Descriptions of log PC_x values, "x" is a selected response like 10% or 50% induction compared to PC_{AGO}, are provided in Figure C.3. log PC_{10} and log PC_{50} values can be defined as the test chemical concentrations estimated to elicit either a 10% or a 50% induction of transcriptional activity induced by PC_{AGO} (AR agonist control; 10 nM of DHT). Each log PC_x value can be calculated by a simple linear regression using two variable data points for the transcriptional activity. Where the data points lying immediately

above and below the log PC_x value have the coordinates (a,b) and (c,d) respectively, then the log PC_x value is calculated using the following equation and Figure C.3:

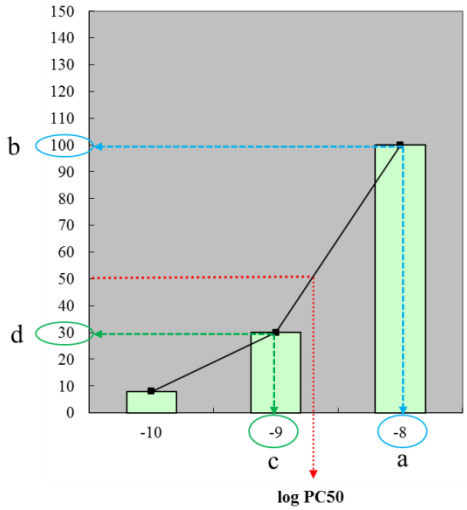
$$\log[PC_x] = c+[(x-d)/(b-d)](a-c)$$

Figure C.3. Schematic descriptions of log PC_x values



The PC_{AGO} (AR agonist control; 10 nM of DHT) is included on each assay plate in agonist assay.
RTA: relative transcriptional activity

Figure C.4. Example for calculation of log PC_{50}

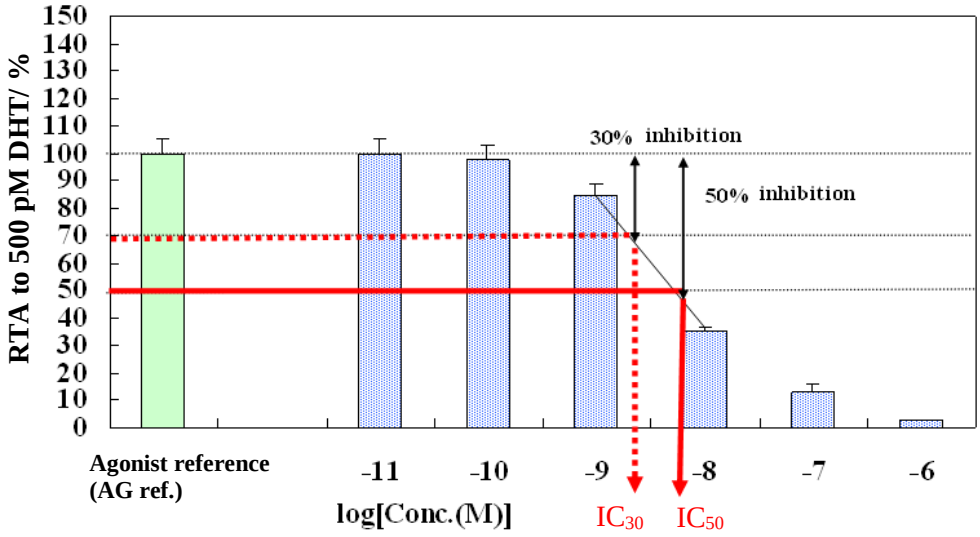


56. In the case of the antagonist assay, the following information should be provided for each positive test chemical: the concentrations of 30% inhibition of transcriptional activity induced by 500 pM DHT (log IC30) and, if appropriate, to 50% inhibition of activity of 500 pM DHT (log IC50).

57. Descriptions of log IC_x values, "x" is a selected response like 30% or 50% inhibition compared to DHT controls, are provided in Figure C.5. log IC_{50} and log IC_{30} values can be defined as the test chemical concentrations estimated to elicit either a 50% or a 30% inhibition of transcriptional activity induced by 500 pM DHT. These values can be calculated in the same way as the log PC values. Each log IC_x value can be calculated by a simple linear regression using two variable data points for the transcriptional activity. Where the data points lying immediately above and below the log IC_x value have the coordinates (c,d) and (a,b) respectively, then the log IC_x value is calculated using the following equation and Figure C.5:

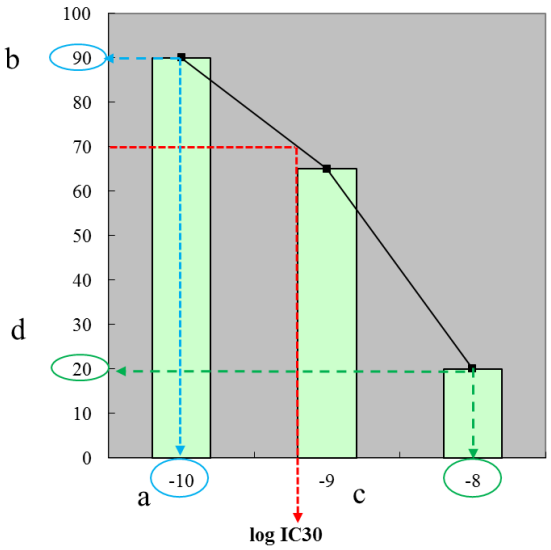
$$\log [IC_x] = a-[(b-(100-x)) /(b-d)] (a-c)$$

Figure C.5. Schematic descriptions of log IC values



The AG ref (DMSO at 0.1% spiked with 500 pM DHT) is included on each assay plate in antagonist assay.
RTA: relative transcriptional activity

Figure C.6. Examples for calculation of log IC_{30}



58. To distinguish pure antagonism from a cytotoxicity-related decrease of luciferase activity, AR-EcoScreen™ is designed to express two kinds of luciferase: firefly luciferase inducibly expressed by the AR response element and renilla luciferase stably and non-inducibly expressed.

59. By using dual reporter assay system, both cell viability and the antagonism can be evaluated in the same cells in a single plate run. The response for the cytotoxic control (10μg/mL of Cycloheximide called PC_{CT}) is used to adjust renilla activity by subtracting the PC_{CT} values – the so-called "renilla activities" - from those of all sample wells. To evaluate the true cytotoxicity of chemicals with the AR Ecoscreen™ assay, such revised

cell viability should be used. If the cell viability is lower than 80% at the specific concentration of a test chemical, this/these data point(s) is/are left out of the calculations.

60. The results, i.e. positive or negative judgment of test chemical, should be based on a minimum of two or three independent runs. If two runs give comparable and reproducible results, it may not be necessary to conduct a third run. To be acceptable, the results should:

- Meet the acceptability criteria (see paragraphs 21-29)
- Be reproducible in triplicate wells (CV<20%).

Data Interpretation Criteria

61. **For the agonist assay,** data interpretation criteria are shown in Table C.4a. Positive results will be characterised by both the magnitude of the effect and the concentration at which the effect occurs. Expressing results as a concentration at which a 50% (log PC_{50}) or 10% (log PC_{10}) are reached accomplishes the goal. However, a test chemical is determined to be positive if the maximum response induction by the test chemical (RPC_{max}) is equal to or exceeds 10% of the reference chemical response in at least two of two or two of three runs, whereas a test chemical is considered negative if the RPC_{max} fails to achieve at least 10% of the reference chemical response in two of two or two of three runs.

Table C.4a. Positive and negative decision criteria for agonist assay

Positive	If a RPC_{max} is obtained that is equal to or exceeds 10% of the response of the positive control.
Negative	If a RPC_{max} fails to achieve at least 10% of the response of the positive control.

62. **For the antagonist assay,** data interpretation criteria are shown in Table C.4b. Positive results will be characterised by both the magnitude of the effect and the concentration at which the effect occurs. Expressing results as a concentration at which a 50% (log IC_{50}) or 30% (log IC_{30}) are reached, accomplishes this goal. However, a test chemical is determined to be positive if the log IC_{30} could be calculated in at least two of two or two of three runs, whereas a test chemical is considered as negative if the log IC_{30} could not be calculated in two of two or two of three runs.

Table C.4b. Positive and negative decision criteria for antagonist assay

Positive	If the log IC_{30} is calculated.
Negative	If the log IC_{30} cannot be calculated.

63. The calculations of log PC_{10}, log PC_{50} and RPC_{max} for agonist assay, and log IC_{50} and log IC_{30} for antagonist assay can be calculated by using a spreadsheet available with the Test Guideline on the OECD public website.

64. It should be sufficient to obtain log PC_x or log IC_x values at least twice. However, should the resulting base-line for data in the same concentration range show variability with high coefficient of variation (% CV), it should be considered that the reliability of the

data is low and the source of the high variability should be identified. The % CV of the raw data triplicate wells (i.e. luminescence intensity data) of the data points on the same assay plate that are used for the calculation of log PC_x or log IC_x should be less than 20%. When an equivocal or inconclusive result is suspected, an additional run or check can be considered.

65. Meeting the acceptability criteria indicates the assay system is operating properly, but it does not ensure that any particular run will produce accurate data. Duplicating the results of the first run is the best assurance that accurate data were produced.

Literature

1. Escande, A., et al. (2006), Evaluation of ligand selectivity using reporter cell lines stably expressing estrogen receptor alpha or beta, *Biochem. Pharmacol.*, 71, 1459-1469.

2. Kuiper, G.G., et al. (1998), Interaction of estrogenic chemicals and phytoestrogens with estrogen receptor beta. *Endocrinology*, 139, 4252-4263.

3. Tarnow, P., Tralau, T., Hunecke, D., Luch, A.(2013), Effects of triclocarban on the transcription of estrogen, androgen and aryl hydrocarbon receptor responsive genes in human breast cancer cells. *Toxicology In Vitro*. 27, 1467-1475.

4. Satoh, K., Nonaka, R., Ohyama, K., Nagai, F. (2005), Androgenic and antiandrogenic effects of alkylphenols and parabens assessed using the reporter gene assay with stably transfected CHO-K1 cells (AR-EcoScreen). *Journal of Health Science*, 51(5), 557–568.

5. Wilson V. S., Bobseine K., Lambright C. R. and Gray Jr. L. E., (2002), A Novel Cell Line, MDA-kb2, That Stably Expresses an Androgen- and Glucocorticoid-Responsive Reporter for the Detection of Hormone Receptor Agonists and Antagonists. *Toxicological. Sciences*. 66 (1): 69-81.

6. OECD (2016), Validation report of Androgen Receptor (AR) Mediated Stably Transfected Transactivation (AR STTA) Assay to Detect Androgenic and Anti-androgenic Activities, Environment, Health and Safety Publications, Series on Testing and Assessment (No.241), Organisation for Economic Cooperation and Development, Paris.

7. Araki N, Ohno K, Nakai M, Takeyoshi M, Iida M. (2005) Screening for androgen receptor activities in 253 industrial chemicals by *in vitro* reporter gene assays using AR-EcoScreen cells. *Toxicology In Vitro.* 19(6):831-842.

8. Araki N, Ohno K, Takeyoshi M, Iida M. (2005) Evaluation of a rapid *in vitro* androgen receptor transcriptional activation assay using AR-EcoScreen cells. *Toxicology In Vitro.* 19(3):335-352.

9. Zwart, N, Andringa, D, de Leeuw, W-J, Kojima, H, Iida, M, Houtman, C J, de Boer, J, Kool, J, Lamoree, M H, Hamers, T. (2017). Improved androgen specificity of AR-EcoScreen by CRISPR based glucocorticoid receptor knockout. *Toxicology in Vitro.* 45, 1–9.

10. Spaepen, M., Angulo, A.F., Marynen, P. and Cassiman, J.J. (1992), Detection of bacterial and mycoplasma contamination in cell cultures by polymerase chain reaction. *FEMS Microbiol Lett.* 78(1), 89-94.

11. Kobayashi, H., Yamamoto, K., Eguchi, M., Kubo, M., Nakagami, S., Wakisaka, S., Kaizuka, M. and Ishii H (1995), Rapid detection of mycoplasma contamination in cell cultures by enzymatic detection of polymerase chain reaction (PCR) products. *J. Vet. Med. Sci.* 57(4), 769-771.

12. Dussurget, O. and Roulland-Dussoix D. (1994), Rapid, sensitive PCR-based detection of mycoplasmas in simulated samples of animal sera. *Appl. Environ. Microbiol*. 60(3), 953-959.

13. De Lean A., Munson P.J. and Rodbard D. (1978). Simultaneous Analysis of Families of Sigmoidal Curves: Application to Bioassay, Radioligand Assay, and Physiological Dose-Response Curves. *American Journal of Physiology*, 235, E97-l02.

Appendix C.1

False positives: Assessment of non-AR-mediated luminescence signals

1. False positives might be generated by non-AR-mediated activation of the luciferase gene, or direct activation of the gene product or unrelated luminescence. Such effects are indicated by an incomplete or unusual dose-response curve. If such effects are suspected, the effect of an AR antagonist (e.g. Hydroxyflutamide (HF) at non-toxic concentration) on the response should be examined.

2. To ensure validity of this approach, the agonistic activity of the following needs to be tested in the same plate:

- Agonistic activity of the chemical with / without 1 µM of HF (in triplicate)
- VC (in triplicate)
- µM HF (in triplicate)
- 500 pM of DHT (in triplicate) as PC_{AGO}

Data interpretation criteria

3. Note: All wells should be treated with the same concentration of the vehicle.

- If the agonistic activity of the chemical is NOT affected by the treatment with HF, it is classified as "Negative".
- If the agonistic activity of the chemical is inhibited, apply the decision criteria (Table C.5a).
- If the agonistic activity at any concentrations tested is inhibited by the treatment with 1 µM of HF (AR antagonist), the difference in the responses between the wells non-treated with the AR antagonist and wells treated with the AR antagonist is calculated. This difference should be considered as the true response and should be used for the calculation of the appropriate parameters to enable a classification decision to be made.

True response = (Response without HF) - (Response with HF)

Data analysis

4. Check the performance standard.

5. Check the CV between wells treated under the same conditions.

a Calculate the mean of the VC
b Subtract the mean of VC from each well value **not** treated with HF
c Calculate the mean of HF
d Subtract the mean of the VC from each well value treated with HF
e Calculate the mean of the PC_{AGO}
f Calculate the relative transcriptional activity of all other wells relative to the PC_{AGO}

Annex D. (Method 2): Androgen Receptor TransActivation Assay for Detection of Androgenic Agonist and Antagonist Activity of Chemicals using the stably transfected human AR-CALUX® cell line[1]

Initial Considerations and Limitations

1. The "General Introduction" should be read before using this test method (Main body page 6-9).

2. The AR-CALUX®(*) transactivation assay uses the human osteosarcoma U2OS AR-CALUX® cell line to detect (anti)androgenic activity mediated through a human androgen receptor (hAR). The AR-CALUX® cell line expresses stably transfected hAR and has no or little expression of other steroid hormone receptors (1).

3. This test method is specifically designed to detect AR-mediated transactivation by measuring bioluminescence as the endpoint. Bioluminescence is commonly used as a read out in various bioassays because of the high signal-to-noise ratio (2). Chemical dependent interference with luminescence signals are reported for certain luc-transformed cell lines but were not observed with the CALUX cell lines.

4. The cell line has low metabolic activity. By combining the test method with a S9 fraction, the impact of metabolism on test chemical activity can be studied (3) and is currently (2020) being validated.

5. The test method has been used for high throughput screening purposes (4). It did not undergo a validation according to the OECD Guidance Document 34.

6. The test method is theoretically applicable to the testing of multi-constituent chemicals and mixtures. During the validation study of this test method single test chemicals were mainly used. When considering testing of mixtures or difficult-to-test chemicals (e.g. unstable) upfront consideration should be given to whether the results of such testing will yield results that are scientifically meaningful.

7. The validation study of the AR-CALUX® test method demonstrated the reliability and relevance of the assay for its intended purpose (5). The test method protocol is described in the referenced document (6).

8. Definitions and abbreviations used in this test method are described in Annex A of this TG.

Principle of the Test Method

9. The test method is used to assess the transactivation of a reporter gene. The AR, when bound to a ligand, is translocated to the nucleus. In the nucleus, the receptor-ligand complex binds specific DNA sequences (androgen responsive elements: AREs) and transactivates a firefly luciferase reporter gene, resulting in increased cellular expression of the luciferase enzyme. Following the addition and subsequent catalytic oxidation of the

[1] Note (*): "CALUX" is a registered trademark, owned by Abraham Brouwer. BioDetection Systems BV (BDS) has obtained the exclusive global right to use and sublicense this trademark.

substrate luciferin, light is emitted. The light produced can easily be detected and quantified using a luminometer.

10. The test system utilises stably transfected AR-CALUX® cells. AR-CALUX® cells originated from the human osteoblastic osteosarcoma U2OS cell line. Human U2OS cells were stably transfected with 3xARE-TATA-Luc and pSG5-neo-hAR using the calcium phosphate co-precipitation method. The U2OS cell line was selected as the good candidate to serve as the androgen - (and other steroid hormone) responsive reporter cell line, based on the observation that the U2OS cell line showed little or no endogenous receptor activity. The absence of endogenous receptors was assessed using luciferase reporter plasmids only, showing no activity when receptor ligands were added. Furthermore, this cell line supported strong hormone-mediated responses when cognate receptors were transiently introduced (1).

11. Testing chemicals for (anti)-androgenic activity using the AR-CALUX® cell line includes a pre-screen run followed by a comprehensive run/specificity control test. During the pre-screen run, the solubility, the cytotoxicity and a refined concentration-range of test chemicals for comprehensive testing are determined. In the subsequent comprehensive run for agonism and antagonism, the test chemical is assessed using the refined concentration-ranges followed by data interpretation. For antagonism, the test chemical is assessed simultaneously with a comprehensive run and a specificity control test.

12. The specificity control test is included to discriminate true competitive antagonists from false positive antagonists (e.g. due to cytotoxicity, cell stress or aspecific inhibition). Cells are exposed to both the EC_{50} concentration and the 100x EC_{50} concentration of the reference agonist DHT when treated with 8 concentrations of the test chemical. This is carried out in the same plate. It will result in two dose responses of which the one generated with the higher ligand concentration (100x EC_{50}) is shifted to the right (see Figure D.6). The shift can be quantitatively measured and an acceptance criterion was developed (R^2). Criteria for data interpretation are described in detail in paragraph 70. Briefly, a test chemical is considered positive for agonism in case at least two consecutive concentrations of the test chemical show a response that is equal or higher than 10% of the maximum response of the reference standard DHT (PC_{10}). A test chemical is considered positive for antagonism in case at least two consecutive concentrations of the test chemical show a response that is equal or lower than 80% of the maximum response of the reference standard Flutamide (PC_{80}) and the specificity control criteria are met.

Demonstration of Laboratory Proficiency

13. Each laboratory should demonstrate proficiency in using this test method prior to testing chemicals with unknown activity. Proficiency is demonstrated by testing the proficiency chemicals for agonist activity and antagonist activity (see Tables B.4a and B.4b in Annex B). This testing will also confirm the responsiveness of the test system. Testing should be replicated at least twice, on different days, and the results should be consistent to the listed classifications and values in Tables B.4a and B.4b. Moreover, an historical database of data generated with the reference standards and the vehicle/solvent controls shall be maintained to confirm the reproducibility of the test method in the respective laboratory over time.

Procedure

Cell line

14. The stably transfected U2OS AR-CALUX® cell line should be used for the test method. The cell line can be obtained from BioDetection Systems BV, Amsterdam, the Netherlands with a technical licensing agreement.

15. Only mycoplasma free cell cultures should be used. Cell batches used should either be certified negative for mycoplasma contamination, or a mycoplasma test should be performed before use. A highly sensitive test, such as PCR methodologies, should be used for detection of mycoplasma infection (8, 9).

Stability of the cell line

16. To maintain the stability and integrity of the AR-CALUX® cells, the cells should be stored < -130°C (e.g. in liquid nitrogen). Following thawing of the cells to start a new culture, cells should be sub-cultured at least twice before being used to assess the (anti)androgenic activity of chemicals. Cells should not be sub-cultured for more than 30 passages.

17. To monitor the stability of the cell line over time, its responsiveness to the reference chemicals (for agonist and antagonist testing) should be verified by evaluating the EC_{50} or IC_{50}. In addition, the relative induction of the positive control (PC) and the negative control (NC) should be monitored. The results should be in agreement with the acceptability criteria for the agonist (Table D.3) or antagonist AR-CALUX® test method (Table D.4). The reference standards, i.e. the reference chemical, positive and negative controls are given in Tables D.1 and D.2 for the agonist and antagonist mode respectively including the concentrations to be used.

Cell culture and plating conditions

18. The AR-CALUX® cells should be cultured in growth medium (DMEM/F12 (1:1)) with phenol red as pH indicator, supplemented with fetal bovine serum (7.5%), non-essential amino acids (1%), penicillin (10 Units/mL), streptomycin (10 µg/mL) and geneticin (G-418) (0.2 mg/mL) as selection marker. Cells should be placed in a CO_2 incubator (5% +/- 1% CO_2) at 37°C +/- 1°C and humidified. When cells reach 85-95% confluency, cells should either be subcultured or prepared for seeding in 96-well microtiter plates. In case of the latter, cells should be resuspended at $1x10^5$ cells/mL in assay medium (DMEM/F12 (1:1)) without phenol red, supplemented with Dextran-Coated Charcoal treated fetal bovine serum (5% v/v), non-essential amino acids (1% v/v), penicillin (10 Units/mL) and streptomycin (10 µg/mL) and plated into 96-well microtiter plates (100 µL of homogenised cell suspension). Cells should be pre-incubated in a CO_2 incubator (5% +/- 1% CO_2, 37°C +/- 1°C, humidified) for 24 hours prior to exposure.

19. Prior to starting any study, all materials (glass tubes, vessels, plastic ware) and reagents (e.g. serum, DMSO) that will be used during the testing should be investigated, as defined in the protocol (6), for any possible interference with the measurements.

Acceptability criteria

20. Agonist and antagonist activities of the test chemical(s) are tested in runs (pre-screen run and comprehensive run). Each run consists of a maximum of 6 microtiter plates.

Each run contains 1 full series of dilutions of a reference chemical (C1 to C8), a fixed concentration of a positive control, a fixed concentration of a negative control, a solvent control (and vehicle control for the antagonist assay) and a positive control for cytotoxicity. In Figures D.1 and D.2, the plate setup for agonist and antagonist runs are given.

21. In the first plate of each run,

- A complete dilution series of the reference chemical (DHT for agonism and Flutamide for antagonism) is measured (Tables D.5 and Table D.6). This reference chemical should demonstrate a sigmoidal dose-response curve. The EC_{50} or IC_{50} derived from the response of the series of dilutions of the reference chemical, and the CV of log(EC_{50}) and log(IC_{50}) for the reference chemicals should fulfil the requirements as indicated in Tables D.3 (agonism) or Table D.4 (antagonism).
- The calculated relative induction of both the positive and negative control should fulfil the requirements as indicated in Tables D.3 and D.4.

22. For each of the microtiter plates within a run, the following is calculated:

- During all measurements, the induction factor of the reference chemical should be calculated by dividing its average relative light unit (RLU) response at the highest concentration (C8) by the average solvent control RLU response. This induction factor should fulfil the minimum requirements for the induction fold as indicated in Tables D.3 and D.4.
- For each test-plate, the Z-factor is calculated according to the equation given below. This Z-factor should fulfil the minimum requirements for the Z-factor as indicated in Tables D.3 and D.4.

$$Z-factor_{plate\ no.} = 1 - 3 * \frac{(SD\ RLU_{plate\ no.}[SC] + SD\ RLU_{plate\ no.}[C8\ reference])}{|(average\ RLU_{plate\ no.}[SC] - average\ RLU_{plate\ no.}[C8\ reference])|}$$

23. A run is considered valid when it fulfils the requirements as stated in Tables D.3 and D.4 and permits to evaluate the response of the test chemicals.

24. The acceptability criteria are applicable to both pre-screen runs and comprehensive runs.

Table D.1. Concentrations of the reference standards for the agonist testing

	Chemical	CASRN	Test range (M) in well
Reference chemical	DHT	521-18-6	1.0×10^{-11} - 1.0×10^{-07}
Positive control (PC)	17α-Methyltestosterone	58-18-4	1.0×10^{-07}
Negative control (NC)	Corticosterone	50-22-6	1.0×10^{-06}

Table D.2. Concentrations of the reference standards for the antagonist testing

	Chemical	CASRN	Test range (M) in well
Reference chemical	Flutamide	13311-84-7	1.0×10^{-08} - 3.0×10^{-05}
Positive control (PC)	Linuron	330-55-2	1.0×10^{-05}
Negative control (NC)	Levonorgestrel	797-63-7	1.0×10^{-06}

Table D.3. Acceptability criteria for agonism pre-screen and comprehensive testing

Number	Acceptability criteria	
1	Sigmoidal curve of reference chemical DHT	Yes
2	EC_{50} range reference chemical DHT	1.0×10^{-10} –1.0×10^{-09} M
3	|CV| of estimated log(EC_{50}) reference chemical DHT	< 1.5%
4	Relative induction (%) PC 17α-Methyltestosterone	> 30%
5	Relative induction (%) NC Corticosterone	< 10%
6	Minimum induction fold of the highest DHT concentration (C8), with respect to the solvent control (SC) on each plate	> 20
7	Z-factor calculated on each plate with DHT C8 and SC	> 0.5

Table D.4. Acceptance criteria for antagonism pre-screen and comprehensive testing

Number	Acceptability criteria	
1	Sigmoidal curve of reference chemical Flutamide	Yes
2	IC_{50} range reference chemical Flutamide	1.0×10^{-07} – 1.0×10^{-06} M
3	|CV| of estimated log(IC_{50}) reference chemical Flutamide	< 3%
4	Relative induction PC (Linuron)	< 60%
5	Relative induction NC (Levonorgestrel)	> 85%
6	Minimum inhibition fold of the highest Flutamide concentration (C8) with respect to the solvent control (SC) on each plate	> 10
7	Z-factor calculated on each plate with Flutamide C8 and SC	> 0.5
8	R^2 between Y_c and S_c^n for Flutamide	≤ 0.7

Solvent/vehicle control and reference standards

25. For both pre-screen and comprehensive runs, the same solvent/vehicle control and the reference standards (reference chemicals, positive controls and negative controls) should be used. In addition, the concentrations of the reference standards should be the same.

Solvent control and vehicle control

26. The solvent used to dissolve test chemicals should solubilize the test chemical completely and should be miscible with the assay medium. DMSO, water and ethanol (95% to 100% purity) are suitable solvents. DMSO (CASRN 67-68-5) is the first choice and its maximum concentration during incubation should not exceed 0.1% (v/v). Prior to the use of another solvent it should be demonstrated that it does not cause cytotoxicity of the cells and interference with the assay performance at exposure concentrations which simulate the experimental conditions.

27. The solvent used to dissolve the test chemicals should also be tested without the dissolved test chemical (solvent control (SC)).

28. For testing agonism, the SC contains assay medium plus the solvent. For testing antagonism, the SC contains the assay medium plus the solvent and a fixed concentration of the agonist reference chemical DHT (the EC_{50} concentration). The vehicle control (VC) however contains the assay medium plus the solvent but does not contain the fixed concentration of the agonist reference chemical.

Reference chemicals

29. The agonist reference chemical is DHT and comprises a series of dilutions of eight concentrations (Tables D.1 and D.5).

30. The antagonist reference chemical is Flutamide and comprises a series of dilutions of eight concentrations (Tables D.2 and D.6). Each of the concentrations of the antagonist reference chemical is spiked with a fixed concentration of the agonist reference chemical

DHT (EC_{50} concentration = 3.0 x 10^{-10} M) in order to measure attenuation of the agonist response.

31. The antagonist reference chemical for the specificity control is Flutamide where each concentration is spiked with a 100X EC_{50} concentration of DHT. It comprises a series of dilutions of eight concentrations (Tables D.2 and D.6).

Positive control

32. The positive control for agonist studies is 17α-Methyltestosterone (Table D.1).

33. The positive control for antagonist studies is Linuron (Table D.2). This control is spiked with a fixed concentration of the agonist reference chemical DHT (3.0 x 10^{-10} M).

Negative control

34. The negative control for agonist studies is Corticosterone (Table D.1).

35. The negative control for antagonist studies is Levonorgestrel (Table D.2). This control is spiked with a fixed concentration of the agonist reference chemical DHT (3.0 x 10^{-10} M).

Preparation of the reference standards and the test chemicals

36. Reference standards (reference chemicals, positive controls, negative controls) and test chemicals are dissolved in 100% DMSO (or an appropriate solvent). Appropriate (serial) dilutions should then be prepared in the same solvent for the reference standards and the test chemicals. Before being dissolved, all chemicals should be allowed to equilibrate to room temperature. Freshly prepared stock solutions of the reference standards and the test chemicals should not have noticeable precipitate or cloudiness.

37. Stock solutions of the reference chemicals (DHT and Flutamide) may be prepared in bulk and stored as aliquots at -20°C +/- 1°C for up to 3 months. Once an aliquot is thawed, it can be stored at -20°C +/- 1°C and re-used (thawing/freezing) for up to 3 weeks. Stock solutions of test chemicals should be prepared fresh before each experiment.

38. Final dilutions of the reference standards and the test chemicals (i.e. working solutions) should be prepared fresh for each experiment and used within 24 hours of preparation.

Solubility, cytotoxicity and range finding

39. Test chemicals shall be assessed at a maximum concentration of 0.1 M (stock solution). When the molecular weight of a test chemical cannot be calculated such as for multi constituent chemicals, polymers, mixtures, UVCBs etc, the gravimetric method should be used starting from 50 mg/mL.

40. The solubility protocol, as used in the validation study, can be found in referenced document (10). Other protocols can be used as long as it is shown that they are suitable by e.g. testing the proficiency chemicals. The solubility of the test chemicals in the solvent of choice should be determined starting from a maximum stock concentration of 0.1 M. In case this concentration shows solubility problems, lower concentrations of stock solutions should be prepared until the test chemicals are fully solubilized. Subsequently, solubility of the test chemical should be assessed in assay medium at exposure concentrations (exposure concentration is 0.1% of the stock concentration, i.e. 0.1 mM).

41. During the pre-screen run, 1:10 serial dilutions of the test chemical are tested. An appropriate refined concentration range for test chemicals is derived from the pre-screen results, to be tested during the comprehensive run. The dilution factor (DF) to be used for comprehensive testing should be as follows: in case a positive response is observed (RI ≥10% for agonism or RI ≤80% for antagonism), an alternating DF 3/3.3 is applied; in case only the highest tested concentration is above the 10% threshold (in agonism testing) or below the 80% threshold (in antagonism testing), DF 2 is applied; in case no response is observed, DF 5 is applied (see Tables D.5 and D.6).

42. Cytotoxicity testing is included in the agonist and antagonist test method protocols, and, is incorporated in both the pre-screen run and comprehensive runs. Following exposure to test chemicals, in pre-screen runs cytotoxicity is assessed with both the lactate dehydrogenase (LDH) leakage test and qualitative visual inspection (i.e. microscopic observation of the cells for morphological changes). The visual inspection is considered as an important evaluation tool given that the LDH leakage test reports on cell death only (cell lysis). For comprehensive runs, the qualitative visual inspection to score cytotoxicity is sufficient. With respect to the LDH leakage test, the concentration of the test chemical is regarded as cytotoxic when the percentage LDH leakage is higher than 15% with respect to the positive control for cytotoxicity (0.01% of Triton X-100). Other cytotoxicity tests can be used as long as it is shown that they are suitable e.g. by testing the proficiency chemicals.

Test chemical exposure and assay plate organisation

43. Following trypsination of a flask of confluent cultured cells, cells are re-suspended in assay medium at $1x10^5$ cells/mL. 100 µL of re-suspended cells are plated in the wells B1-G11 of a 96-well microtiter plate. The remaining wells are filled with 200 µL of Phosphate Buffered Saline (PBS) (see Figures D.1 and D.2). The plated cells are pre-incubated for 24 +/- 8 hours in a CO_2 incubator (5% +/- 1% CO_2, 37°C+/- 1°C, humidified).

44. After pre-incubation, the condition of the cells is verified visually (cytotoxicity, contamination and confluence (microscopy)). Only plates that show no visual cytotoxicity or contamination and have a minimum of 85% confluence in a representative portion of all wells are used for testing. Cells in the wells B1-G11 are exposed by the addition of 100 µL of assay medium containing appropriate dilution series of the reference standards, the test chemicals, the solvent controls and cytotoxicity controls (Table D.5: agonist testing; Table D.6: antagonist testing).

45. All reference standards, test chemicals and solvent controls are tested in triplicate whereas the cytotoxicity control Triton X-100 is tested in six replicate wells. In Figure D.1, the plate layout for agonist testing is given which is identical for pre-screen testing and comprehensive testing. In Figure D.2, the plate layout for antagonist pre-screen testing is given. All exposed wells, except for the vehicle control wells (VC), contain a fixed concentration of agonist reference chemical DHT (3.0×10^{-10} M (EC_{50})). In Figure D.3, the plate layout for antagonist comprehensive testing is shown, including the specificity control test with 100x EC_{50} of DHT. This corresponds to C(1-8)100 in the plate layout.

46. The 96-well microtiter plates should be incubated for another 24 +/- 2 hours in a CO_2 incubator (5% +/- 1% CO_2, 37°C +/-1°C, humidified). After incubation, the plates are visually inspected for cytotoxicity and contamination. For pre-screen testing, 100 µl of the exposure medium from each well is transferred to another plate to be used for the cytotoxicity test (paragraph 42). The remaining 100 µl of exposure medium in the well is

removed in order to expose the cells in the wells to lysis substrate (paragraph 47) for measuring luminescence.

Figure D.1. Plate layout of the 96-well microtiter plates for agonist **pre-screen and comprehensive testing and for assessment of <u>agonist</u> effects.**

Plate 1

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B	Triton X-100	SC	C1 DHT	C2 DHT	C3 DHT	C4 DHT	C5 DHT	C6 DHT	C7 DHT	C8 DHT	PC	
C	Triton X-100	SC	C1 DHT	C2 DHT	C3 DHT	C4 DHT	C5 DHT	C6 DHT	C7 DHT	C8 DHT	PC	
D	Triton X-100	SC	C1 DHT	C2 DHT	C3 DHT	C4 DHT	C5 DHT	C6 DHT	C7 DHT	C8 DHT	PC	
E	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	NC	
F	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	NC	
G	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	NC	
H												

Subsequent plates

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	C8 DHT	
C	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	C8 DHT	
D	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	C8 DHT	
E	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	C8 DHT	
F	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	C8 DHT	
G	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	C8 DHT	
H												

C(1-8) DHT = series of dilutions (1-8, low-to-high concentrations) of reference chemical DHT
C(1-8) = series of dilutions (1-8, low-to-high concentrations) of test chemical.
SC = solvent control of the test chemical / reference standards (the same solvent as in C (1-8)).
PC = positive control 17α-Methyltestosterone.
NC = negative control Corticosterone
Grey cells: = outer wells, filled up with 200 µL of PBS.
Triton X-100 = positive control for cytotoxicity

Figure D.2. Plate layout of the 96-well microtiter plates for antagonist **pre-screen testing and assessment of <u>antagonist</u> effects.**

Plate 1

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B	Triton X-100	SC	C1 FLU	C2 FLU	C3 FLU	C4 FLU	C5 FLU	C6 FLU	C7 FLU	C8 FLU	VC	
C	Triton X-100	SC	C1 FLU	C2 FLU	C3 FLU	C4 FLU	C5 FLU	C6 FLU	C7 FLU	C8 FLU	VC	
D	Triton X-100	SC	C1 FLU	C2 FLU	C3 FLU	C4 FLU	C5 FLU	C6 FLU	C7 FLU	C8 FLU	VC	
E	Triton X-100	NC	C1	C2	C3	C4	C5	C6	C7	C8	PC	
F	Triton X-100	NC	C1	C2	C3	C4	C5	C6	C7	C8	PC	
G	Triton X-100	NC	C1	C2	C3	C4	C5	C6	C7	C8	PC	
H												

Subsequent plates

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	C8 FLU	
C	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	C8 FLU	
D	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	C8 FLU	
E	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	C8 FLU	
F	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	C8 FLU	
G	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	C8 FLU	
H												

C(1-8) FLU = series of dilutions (1-8, low-to-high concentrations) of reference chemical Flutamide (spiked with EC_{50} concentration DHT)
C(1-8) = series of dilutions (1-8, low-to-high concentrations) of test chemical (spiked with EC_{50} concentration DHT)
NC = negative control Levonorgestrel (spiked with EC_{50}concentration DHT)

PC = positive control Linuron (spiked with EC_{50} concentration DHT)
SC = solvent control of the test chemical/ reference standards (the same solvent as in C (1-8) (spiked with EC_{50}concentration DHT).
VC = vehicle control (solvent control without the spiking with DHT).
Grey cells = outer wells, filled up with 200 µL of PBS.
Triton X-100 = positive control for cytotoxicity

Figure D.3. Plate layout of the of the 96-well microtiter plates for antagonist comprehensive testing and assessment of antagonist effects, including specificity control test.

Plate 1

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B	Triton X-100	SC	C1 FLU	C2 FLU	C3 FLU	C4 FLU	C5 FLU	C6 FLU	C7 FLU	C8 FLU	VC	
C	Triton X-100	SC	C1 FLU	C2 FLU	C3 FLU	C4 FLU	C5 FLU	C6 FLU	C7 FLU	C8 FLU	VC	
D	Triton X-100	SC	C1 FLU	C2 FLU	C3 FLU	C4 FLU	C5 FLU	C6 FLU	C7 FLU	C8 FLU	VC	
E	Triton X-100	NC	C1 FLU 100	C2 FLU 100	C3 FLU 100	C4 FLU 100	C5 FLU 100	C6 FLU 100	C7 FLU 100	C8 FLU 100	PC	
F	Triton X-100	NC	C1 FLU 100	C2 FLU 100	C3 FLU 100	C4 FLU 100	C5 FLU 100	C6 FLU 100	C7 FLU 100	C8 FLU 100	PC	
G	Triton X-100	NC	C1 FLU 100	C2 FLU 100	C3 FLU 100	C4 FLU 100	C5 FLU 100	C6 FLU 100	C7 FLU 100	C8 FLU 100	PC	
H												

Subsequent plates

	1	2	3	4	5	6	7	8	9	10	11	12
A												
B	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	C8 FLU	
C	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	C8 FLU	
D	Triton X-100	SC	C1	C2	C3	C4	C5	C6	C7	C8	C8 FLU	
E	Triton X-100	SC	C1 100	C2 100	C3 100	C4 100	C5 100	C6 100	C7 100	C8 100	C8 FLU	
F	Triton X-100	SC	C1 100	C2 100	C3 100	C4 100	C5 100	C6 100	C7 100	C8 100	C8 FLU	
G	Triton X-100	SC	C1 100	C2 100	C3 100	C4 100	C5 100	C6 100	C7 100	C8 100	C8 FLU	
H												

C(1-8) FLU =series of dilutions (1-8, low-to-high concentrations) of reference chemical Flutamide (spiked with EC_{50} concentration DHT)
C(1-8) FLU 100 =series of dilutions (1-8, low-to-high concentrations) of reference chemical Flutamide (spiked with 100X EC_{50} concentration DHT)
NC = negative control Levonorgestrel (spiked with EC_{50} concentration DHT)
PC = positive control Linuron (spiked with EC_{50} concentration DHT)
SC = solvent control of the test chemical/ reference standards (the same solvent as in C (1-8) (spiked with EC_{50} concentration DHT).
VC = vehicle control (solvent control without the spiking with DHT).
C(1-8) = series of dilutions (1-8, low-to-high concentrations) of test chemical (spiked with EC_{50} concentration DHT)
C(1-8) 100 = series of dilutions (1-8, low-to-high concentrations) of test chemical ((spiked with 100x EC_{50} concentration DHT) (specificity test)
Grey cells = outer wells, filled up with 200 µL of PBS.
Triton X-100 = positive control for cytotoxicity

Measurement of luminescence

47. There are several options for the measurement of luminescence. The methods used during the validation of the AR-CALUX® test method included either the use of a commercial kit, which could be either a flash or a glow luminescence kit, or the preparation of the luminescence substrate in-house. In any case, the medium from the wells should be removed and the cells should be lysed following 24 +/- 2 hours of incubation to measure luciferase activity.

48. For measuring the luminescence, a luminometer is required. When transparent plates are used, the luminometer has to be equipped with 2 injectors. The luciferase reaction is started by injection of the substrate luciferin. The reaction is stopped by addition of an appropriate solvent (e.g. 0.2 M NaOH or 25% v/v acetic acid depending on the luminometer) to prevent carry over of luminescence from one well to the other. When a

commercial kit is used, the specific instructions supplied with the kit must be followed. White plates allow the use of a luminometer without or with one injector depending on the kit used.

49. Light emitted from each well is expressed as Relative Light Units (RLUs) per well.

Pre-screen run for (ant)agonist testing

50. The pre-screen analysis results are used to determine a refined concentration-range of the test chemicals for the comprehensive testing. Evaluation of pre-screen analysis results and the determination of the refined concentration-range of test chemicals for comprehensive testing is described in depth in the agonist and antagonist test method protocol (6). Here, a brief summary of the procedures for determining the concentration range of the test chemicals for agonist and antagonist testing is given. See Tables D.5 and D.6 for guidance of serial dilution design.

51. During the pre-screen run, test chemicals should be tested using the dilution series as indicated in Tables D.5 (agonism) and D.6 (antagonism). All concentrations should be tested in triplicate wells according to the plate layout as indicated in Figures D.1 (agonism) or D.2 (antagonism).

52. A pre-screen run shall always be followed by a comprehensive run, regardless if the response observed during the pre-screen is positive or negative. One comprehensive run shall suffice for drawing a conclusion following the decision criteria in Table D.7.

Selection of concentrations for assessment of (ant)agonist effects

53. Only results that fulfil the acceptability criteria (Tables D.3 and D.4) are considered valid and allow evaluating the response to test chemicals. In case one or more microtiter plates in a run fail to fulfil the acceptability criteria, the respective microtiter plates should be tested again. In case the first plate containing the complete series of dilutions of the reference chemical fails the acceptability criteria, the complete run (6 plates) has to be tested again.

54. Determine the (lowest) concentration at which maximum induction (agonism) or inhibition (antagonism) is observed and does not show cytotoxicity. The highest concentration of the test chemical to be tested in the comprehensive run should be 3-times this selected concentration or a maximum exposure concentration of 0.1 mM or 50 µg/ml for chemicals where the molarity is not known.

55. A complete refined dilution series of the test chemical should be prepared with dilutions steps as indicated in Tables D.5 and D.6, starting with the highest concentration as determined above

56. A test chemical that does not elicit any (ant)agonist effect, should be tested in the comprehensive run starting with the highest, non-cytotoxic concentration identified during the pre-screen with dilutions steps as indicated in Tables D.5 and D.6.

Comprehensive run for agonist testing

57. Following the selection of the refined concentration ranges, test chemicals should be tested comprehensively using the dilution series indicated in Table D.5 (agonism). All concentrations should be tested in triplicate wells according to the plate layout as indicated in Figure D.1 (agonism) (see paragraph 45).

58. Only results that fulfil the acceptability criteria (Table D.3) are considered valid and allow evaluating the response to test chemicals. In case one or more microtiter plates in a run fail to fulfil the acceptability criteria, the chemicals tested in the respective microtiter plates should be tested again in a repeated run. In case the first plate containing the complete series of dilutions of the reference chemical fails the acceptability criteria, the complete run (6 plates) has to be tested again.

Comprehensive run and specificity control test for antagonist testing

59. Following the selection of the refined concentration ranges, a test chemical shall be tested simultaneously with a comprehensive test and a specificity control test (on the same plate), using the dilutions series indicated in Table D.6 (antagonism) (see paragraph 45). All concentrations should be tested in triplicate wells according to the plate layout as indicated in Figure D.3 (antagonism).

60. Only results that fulfil the acceptability criteria (Table D.4) are considered valid and allow evaluation of the response of the test chemicals. In case one or more microtiter plates in an analysis series fail to fulfil the acceptance criteria, the respective microtiter plates should be tested again. In case the first plate containing the complete series of dilutions of the reference chemicals fails the acceptability criteria, the complete run (6 plates) has to be tested again.

Table D.5. Concentration and dilutions of reference standards and test chemicals used for agonist testing

Reference DHT conc. (M) in well		Controls conc. (M) in well		Test chemical	Pre-screen	Comprehensive		
					DF 10	DF 5	DF 3/ 3.33	DF 2
C1	1.0×10^{-11}	PC	1.0×10^{-07}	C1	10,000,000 x	78125 x	3,000 x	128 x
C2	3.0×10^{-11}	NC	1.0×10^{-06}	C2	1,000,000 x	15625 x	1,000 x	64 x
C3	1.0×10^{-10}	SC	0	C3	100,000 x	3125 x	300 x	32 x
C4	3.0×10^{-10}			C4	10,000 x	625 x	100 x	16 x
C5	1.0×10^{-09}			C5	1,000 x	125 x	30 x	8 x
C6	3.0×10^{-09}			C6	100 x	25 x	10 x	4 x
C7	1.0×10^{-08}			C7	10 x	5 x	3 x	2 x
C8	$1.0 \times *10^{-07}$			C8	1 x	1 x	1 x	1 x

PC - positive control (17α-Methyltestosterone)
NC - negative control (Corticosterone)
SC - test chemical solvent control

Table D.6. Concentration and dilutions of reference standards and test chemicals used for antagonist testing

Reference Flutamide conc. (M) in well		Controls conc. (M) in well		Test chemical	Pre-screen DF 10	Comprehensive DF 5	DF 3/3.33	DF 2
C1	1.0 x 10^{-08}	PC	1.0 x 10^{-05}	C1	10,000,000 x	78125 x	3,000 x	128 x
C2	3.0 x 10^{-08}	NC	1.0 x 10^{-06}	C2	1,000,000 x	15625 x	1,000 x	64 x
C3	1.0 x 10^{-07}	SC	0	C3	100,000 x	3125 x	300 x	32 x
C4	3.0 x 10^{-07}	VC	0	C4	10,000 x	625 x	100 x	16 x
C5	1.0 x 10^{-06}			C5	1,000 x	125 x	30 x	8 x
C6	3.0 x 10^{-06}			C6	100 x	25 x	10 x	4 x
C7	1.0 x 10^{-05}			C7	10 x	5 x	3 x	2 x
C8	3.0 x 10^{-05}			C8	1 x	1 x	1 x	1 x
		Supplemented agonist conc. (M in well)						
		DHT 3.0 x 10^{-10} (=EC_{50})						
		Supplemented agonist conc. Specificity control (M in well)						
		DHT 3.0 x 10^{-8}(=100x EC_{50})						

PC - positive control (Linuron)
NC - negative control (Levonorgestrel)
SC - solvent control
VC - vehicle control (does not contain fixed concentration of the agonist reference chemical)l

Analysis of Data

Normalisation of the data

61. Raw data derived from the luminometer are expressed as RLUs. When the acceptability criteria are met, as indicated in Tables D.3 and D.4, the following calculation steps are performed to determine the required parameters. The raw data should be transferred to a data analysis spreadsheet designed for pre-screen or comprehensive runs.

For the agonist assay:

62. For each test chemical and the reference chemical DHT calculate

- the relative induction at concentration *c*, technical replicate i , (Y_{ic}) as follows:

$$Y_{ic} = \frac{RLU\ of\ test\ chemical\ (i\ replicate) - average\ RLU\ of\ SC}{average\ RLU\ of\ DHT_{C8} - average\ RLU\ of\ SC} \times 100, \quad i = 1,2,3$$

- the average of the relative inductions over the 3 technical replicates (Y_c),

$$Y_c = average\ Y_{ic}$$

For the antagonist assay:

63. For each test chemical, and reference chemical FLU calculate

- the relative induction at concentration *c*, technical replicate *i*, (Y_{ic}) as follows: .

$$Y_{ic} = \frac{RLU\ of\ test\ chemical\ (i\ replicate) - average\ RLU\ of\ FLU_{C8}}{average\ RLU\ of\ SC - average\ RLU\ of\ FLU_{C8}} \times 100, \quad i = 1,2,3$$

- the average of the relative inductions over the 3 technical replicates (Y_c)

$$Y_c = average\ Y_{ic}$$

Note: SC in the antagonist assay is assay medium spiked with DHT EC_{50} concentration

For the specificity control test

- To calculate the test chemical's specificity control at concentration *c*, replicate *i*, S_{ic}, apply the same formula as given above for the antagonist assay but the RLU of test chemical (i replicates) shall be those obtained in the specificity control test (spiking with 100x EC_{50} DHT).
- In addition calculate the test chemical's normalized specificity control (S_c^n) by setting the C1 concentration of the test chemical's specificity control (S_c) at 100%, i.e.

$$S_c^n = 100 \times \frac{S_c}{S_{c_1}}, \qquad c = c_1, \ldots, c_8$$

Cytotoxicity

64. For all test chemical concentrations evaluated during the pre-screen analysis, calculate the percentage LDH leakage with respect to the cytotoxicity positive control 0.01% Triton X-100 (percentage set at 100%), according to the following equation:

$$\%\ LDH\ leakage = \frac{average\ AU\ test\ chemical\ - average\ AU\ SC}{average\ AU\ positive\ control\ - average\ AU\ SC} \times 100$$

Note: AU= absorbance unit

65. In addition, qualitative visual inspection of the cells following exposure to the test chemicals shall be carried out. The test chemical is regarded cytotoxic at a specific concentration when:

- either the average percentage LDH leakage of the triplicate sample is higher than 15% with respect to the positive control
- or, cytotoxicity is observed with a microscope.

Calculation of parameters

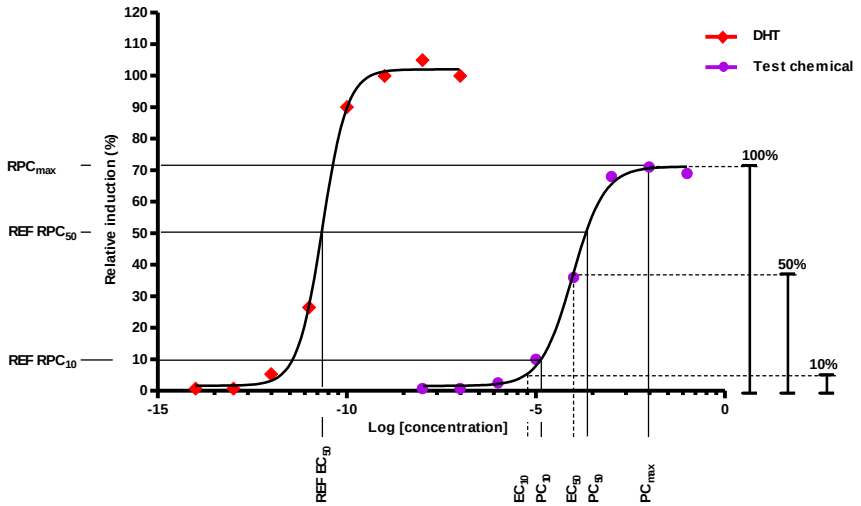
66. After the normalization of the data, apply a non-linear regression (variable slope, 4 parameters) to the Y_{ic} data using the following equation:

$$y = Bottom + \frac{(Top - Bottom)}{\left(1 + 10^{((LogEC_{50} - x) * HillSlope)}\right)}$$

x =	Log of dose or concentration
y =	Response (relative induction (%))
Top =	Maximum induction (%)
Bottom =	Minimum induction (%)
$LogEC_{50}$ =	Log of concentration at which 50% of maximum response is observed
HillSlope =	Slope factor of Hill slope

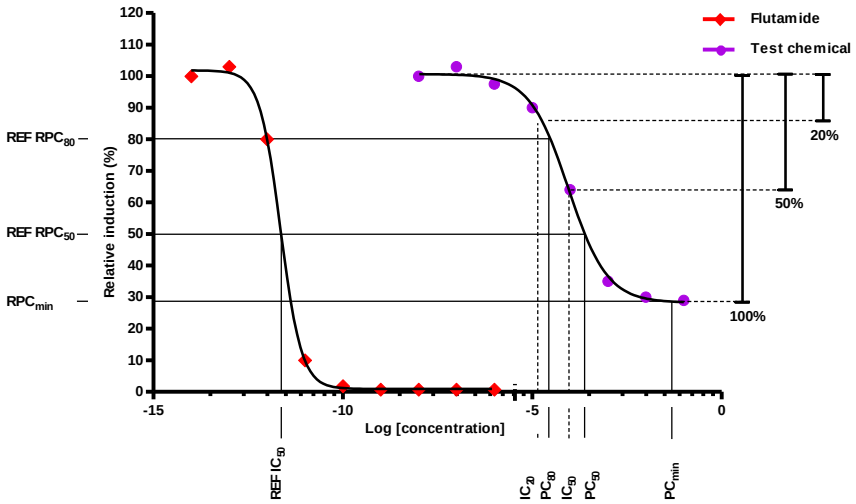
67. For agonist testing, determine the EC_{10} and EC_{50} of the reference chemical, and, determine the EC_{10}, EC_{50}, PC_{10}, and PC_{50} of the test chemicals. For antagonist testing, determine IC_{50} and IC_{20} of the reference chemical, and, determine the IC_{20}, IC_{50}, PC_{80}, and PC_{50} of the test chemicals. To further characterise the potency of a test chemical, the magnitude of the effect (agonism: RPC_{max}; antagonism: RPC_{min}) and the concentration at which the effect occurs (agonism: PC_{max}; antagonism: PC_{min}) should be reported. In Figures D.4 (agonism) and D.5 (antagonism), a graphical representation of these parameters are given.

Figure D.4. Overview of parameters determined for a test chemical in the agonist assay



EC_{10} = concentration of a test chemical at which 10% of its maximum response is observed.
EC_{50} = concentration of a test chemical at which 50% of its maximum response is observed.
PC_{10} = concentration of a test chemical at which its response is equal to the EC_{10} of the reference chemical (REF RPC_{10}).
PC_{50} = concentration of a test chemical at which its response is equal to the EC_{50} of the reference chemical (REF RPC_{50}).
PC_{max} = concentration of a test chemical where the response is maximal (corresponding to RPC_{max})
REF EC_{50} = concentration of the reference chemical DHT at which 50% of its maximum response is observed

Figure D.5. Overview of parameters determined for a test chemical in the antagonist assay.



IC_{20} = concentration of a test chemical at which 80% of its maximum response is observed (20% inhibition).

IC_{50} = concentration of a test chemical at which 50% of its maximum response is observed (50% inhibition).

PC_{80} = concentration of a test chemical at which its response is equal to the IC_{20} of the reference chemical (REF RPC_{80}).

PC_{50} = concentration of a test chemical at which its response is equal to the IC_{50} of the reference chemical (REF RPC_{50})

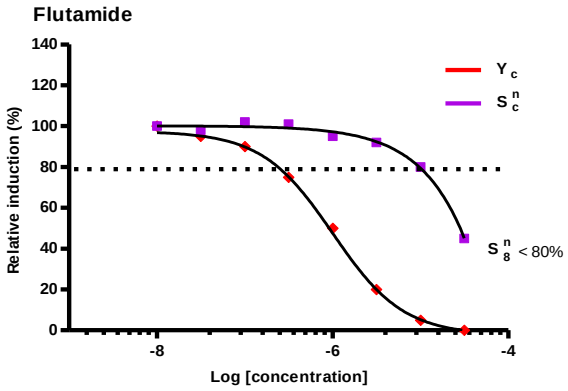
PC_{min} = concentration of a test chemical where the response is maximal (corresponding to RPC_{min})

REF IC_{50} = concentration of the reference chemical at 50% of its maximum response

68. For test chemicals, a full dose-response curve may not always be achieved due to e.g. cytotoxicity or solubility problems. In these cases, the EC_{50} and EC_{10} cannot be determined in the agonist testing, and, the IC_{50} and IC_{20} cannot be determined in the antagonist testing. Therefore it is then sufficient to determine the PC_{10}, PC_{50} and PC_{max} (agonist) and the PC_{80}, PC_{50} and PC_{min} (antagonist), if possible, which can be derived from e.g. linear interpolation between the two closest data points.

69. Specificity of an antagonist response (i.e. being a true competitive antagonist) is determined as indicated by the data interpretation criteria (Table D.7). When interpreting the results of the specificity control, the two dose response curves (Y_c and S_c^n) should be visually inspected and it should be verified whether the first positive criterion for antagonist testing can be applied (see Table D.7: S_c^n> 80% at all concentrations). Otherwise, calculate the square of the correlation coefficient (R^2) between the relative induction of the standard response (Y_c) and the relative induction of the normalized specificity response (S_c^n) of a test chemical. This second positive criterion of the antagonist testing should be verified (see Table D.7: R^2 is ≤ 0.9). Some caution should be applied as this criterion cannot be considered as 100% definitive (as shown in the AR-CALUX® validation study (5)). It may be influenced by the shape of the curves and by outliers. Expert judgment may need to be applied.

Figure D.6. Representation of the relative induction of the standard response (Y_c) and the relative induction of the normalized specificity response (S_c^n) of a true competitive antagonist ($R^2 \leq 0.9$)



Data Interpretation Criteria

70. For the interpretation of data and the decision whether a test chemical is considered positive or negative, the criteria in Table D.7 are to be used. One comprehensive run shall suffice for drawing a conclusion.

Table D.7. Decision criteria

AGONISM	
Positive	When the relative induction (Y_c) of the test chemicalis ≥ 10% (REF RPC_{10}) for two or more consecutive concentrations.
Negative	In all other cases
ANTAGONISM	
Positive	When the relative induction (Y_c) of the test chemical is ≤ 80% (REF RPC_{80}) for two or more consecutive concentrations and Either • the relative induction of the test chemical's normalised specificity control s_c^n > 80% at all concentrations or when the following two conditions are met: • the relative induction of the test chemical's normalised specificity control at the highest concentration $s_{c_8}^n$ is ≤ 80%, • the square of the correlation coefficient between the relative induction of the test chemical's normalized specificity control (s_c^n) and the relative induction (Y_c) (R^2) is ≤ 0.9
Negative	In all other cases

71. The given criteria in Table D.7 should be applied only to data that were generated in the absence of cytotoxicity.

72. In addition to the dichotomous categorisation (Table D.7), the potency measurements may also be used in integrated approaches.

Literature

1. Sonneveld, E., Jansen, H.J., Riteco, J.A., Brouwer, A, van der Burg B. (2004). Development of androgen- and estrogen-responsive bioassays, members of a panel of human cell line-based highly selective steroid-responsive bioassays. Toxicol. Sci. 83(1), 136-148.
2. Thorne, N., Inglese, J. and Auld, D.S. (2010). Illuminating Insights into Firefly Luciferase and Other Bioluminescent Reporters Used in Chemical Biolog. *Chemistry and Biology* 17(6), 646-57.
3. van Vugt-Lussenburg, B.M.A., van der Lee, R.B., Man, H.Y., Middelhof, I., Brouwer, A., Besselink, H., van der Burg, B. (2018) Incorporation of metabolic enzymes to improve predictivity of reporter gene assay results for estrogenic and anti-androgenic activity. *Reprod Toxicol* 75, 40-48
4. Van der Burg, B., Pieterse, B., Buist, H., Lewin, G., van der Linden, S.C., Man, H.Y., Rorije, E., Piersma, A.H., Mangelsdorf, I., Wolterbeek, A.P., Kroese, E.D., van Vught-Lussemburg, B. (2015a). A high throughput screening system for predicting chemically-induced reproductive organ deformities. *Reprod. Toxicol*. 55, 95-103.
5. Validation Study Report on the Performance assessment of the AR-CALUX® *in vitro* method (2019). Available at (https://tsar.jrc.ec.europa.eu/test-method/tm2010-07)
6. Transactivation assay for the detection of compounds with (anti)androgenic potential using AR-CALUX® cells (2019). Available at (https://tsar.jrc.ec.europa.eu/test-method/tm2010-07)
7. Kleinstreuer, N.C., Ceger, P., Watt, E.D., Martin, M., Houck, K., Browne, P., Thomas, R.S., Casey, W.M., Dix, D.J., Allen, D., Sakamuru, S., Xia, M., Huang, R., Judson, R. (2017), Development and Validation of a Computational Model for Androgen Receptor Activity. *Chem Res Toxicol*., 30(4):946-964
8. Kobayashi, H., Yamamoto, K., Eguchi, M., Kubo, M., Nakagami, S., Wakisaka, S., Kaizuka, M. and Ishii, H. (1995). Rapid detection of mycoplasma contamination in cell cultures by enzymatic detection of polymerase chain reaction (PCR) products. *J. Vet. Med. Sci*., 57(4), 769-771.
9. Eldering, J.A., Felten, C., Veilleux, C.A. & Potts, B.J. Development of a PCR method for mycoplasma testing of Chinese hamster ovary cell cultures used in the manufacture of recombinant therapeutic proteins. Biologicals 32, 183-93 (2004).
10. Solubility Determination by Visual Inspection (2019). Available at (https://tsar.jrc.ec.europa.eu/test-method/tm2010-07)

Annex E. (Method 3): Androgen Receptor Transactivation Assay for Detection of androgenic Agonist and Antagonist Activity of Chemicals using the stably transfected human 22Rv1/MMTV_GR-KO cell line

Initial Considerations and Limitations

1. The "General Introduction" should be read before using this test method (Main body page 6-9).

2. The 22Rv1/MMTV_GR-KO AR-mediated stably transfected transcriptional activation (TA) assay was established to screen chemicals for endocrine activity via interaction with the AR using a human prostate cancer cell line, 22Rv1, that endogenously expresses the AR (1, 2). This test method is specifically designed to detect human AR-mediated TA and inhibition by measuring luciferase activity as the endpoint. The information of chemical dependent interference with luminescence signals is limited in a GR-knockout 22Rv1/MMTV cell line.

3. Although the constitutively-acting truncated AR is expressed in 22Rv1/MMTV cells, the truncated AR does not significantly affect the activity. It is verified that the solvent control level (basal level) is not high, the induction fold is dose-dependently increasing by treatment with DHT, and the level of increase is very high compared to other reporter gene assay (2, 4). Furthermore, the full length AR is expressed to similar level with LNCaP cell (2).

4. The glucocorticoid receptor (GR) is expressed in 22Rv1 origin cells alongside AR, endogenously. The minimal GR-mediated response can interfere with the AR-mediated response because the GR is structurally similar to the AR and shares hormone response elements that exhibit cross-talk with the AR (1, 3, 5). To eliminate GR expression in cells, a GR-knockout 22Rv1/MMTV cell line was developed using the clustered regularly interspaced short palindromic repeats (CRISPR)/CRISPR-associated (Cas) 9 system (1, 2, 5).

5. The GR-knockout 22Rv1/MMTV cell line showed low metabolic activity in validation study (6). The validation was conducted using only monoconstituent chemicals. This test method can theoretically be applied to the testing of mixtures. Before applying this test method to mixtures, it should be considered whether the results will be scientific meaningful.

6. Definitions and abbreviations used in this test guideline are described in Annex A of this TG.

7. The 22Rv1/MMTV_GR-KO assay was validated by the National Institute for Food and Drug Safety Evaluation (NIFDS), the Korean Testing and Research Institute (KTR) and Dongguk University with support of a study management team comprised of members of the OECD VMG-NA expert group. The test method is used to detect AR agonists and antagonists of level 2 in "OECD Conceptual Framework for the Testing and Assessment of Endocrine Disrupting Chemicals" (6, 7, 8). The validation study of the 22Rv1/MMTV_GR-KO method was conducted according to OECD Guidance Document (GD) 34. The relevance and reliability of the assay for its intended purpose was demonstrated (9, 10).

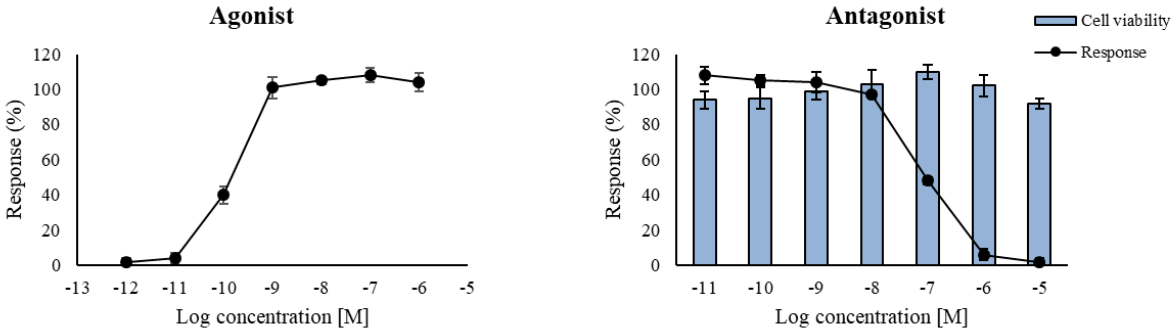
Principle of the test method

8. The test system provided in this method utilises the 22Rv1/MMTV_GR-KO cell line, which is derived from a 22Rv1 cell line. The 22Rv1 cells have been classified as a biosafety level 2 cell line from ATCC, because the 22Rv1 cell line produces the human retrovirus XMRV (xenotropic murine leukemia virus-related virus) (11). When conducting the experiment using the 22Rv1 cell line, the biological safety should be considered. The cell line developed consists of stably transformed 22Rv1 cells with one pGL4[*luc2P*/MMTV/Hygro] vector. The pGL4[*luc2P*/MMTV/Hygro] vector protocol is subject to a Promega limited use licence requiring i) the use of luminescent assay reagents purchased from Promega; or ii) to contact Promega to obtain a free license for commercial use.

9. AR agonist/antagonist assays using the 22Rv1/MMTV_GR-KO cell line should be conducted in a stepwise approach. After conduct of a pre-screen run, a comprehensive run and specificity control (only AR antagonist assay) are performed. The comprehensive run is only conducted if the pre-screen indicates positive activity in either the agonist or antagonist assay. A starting concentration of the test chemical for a comprehensive run is determined in the pre-screen run. To confirm whether a chemical is a true competitive AR binding antagonist, a specificity control test must be used (see paragraph 28).

10. Data interpretation for an AR agonistic effect is based upon the maximum response level induced by a test chemical. If this response equals or exceeds 10% of the response induced by 10 nM 5α-DHT, the AR agonist control (PC_{AGO}), the test chemical is considered a AR agonist. Data interpretation for an AR antagonist effect of a test chemical is decided by two steps. i) a cut-off of 30% inhibitory response of the test chemical in the presence of 800 pM DHT and ii) R^2 value less than 0.9 in the specificity control test (see paragraphs 40 and 43). If both criteria are met, then the chemical is considered a true AR antagonist. Data analyses is described in detail in paragraphs 37-41. Typical concentration-response curves of agonist and antagonist reference chemical (DHT and Bicalutamide) are shown in Figure E.1.

Figure E.1. Typical positive control responses from the pre-screen run in AR agonist and antagonist assay



Demonstration of laboratory proficiency

11. A proficiency test should be conducted by each laboratory to verify proficiency with the 22Rv1/MMTV GR-KO method. The proficiency chemicals for the agonist and antagonist assay are listed in Tables B.4a and B.4b of the Annex B in this TG. The proficiency test should be done at least twice, on different days, and the results should be consistent with the classifications and values for the proficiency chemicals listed in Tables B.4a and B.4b in Annex B.

Procedure

Cell line

12. The 22Rv1/MMTV_GR-KO cell line is an androgen-responsive stable transformed cell line derived from 22Rv1 human prostate cancer cells, which are adherent and AR-positive. The cell line can be obtained from the Korean Cell Bank, upon signing a Material Transfer Agreement (MTA) containing a license agreement.

13. Mycoplasma-free cells should be used in the method. The detection of mycoplasma infection should be conducted before starting any experiments using sensitive methods, such as PCR analysis (12).

Stability of the cell line

14. To maintain the stability and integrity of the response, the cells should be kept at less than -80°C (e.g. in deep freezer or liquid nitrogen). Cells should be sub-cultured at least twice after thawing, and shall than be used to assess the (anti)androgenic activity of chemicals. Cells should not be sub-cultured for more than 30 passages The cell-doubling time is 48 hours.

Cell line maintenance and plating conditions

15. The following medium should be prepared (the details are described in the SOP of the validation report (10)):

- ▪ Culture medium: RPMI1640 supplemented with FBS (10% v/v), GlutaMAX™ (2 mM), Penicillin (100 units/mL), Streptomycin (100 µg/mL), and Amphotericin B (0.25 µg/mL).
- ▪ Test medium: phenol red-free RPMI1640 supplemented with Dextran-coated charcoal treated (DCC)-FBS (5% v/v), GlutaMAX™ (2 mM), Penicillin (100 units/mL), Streptomycin (100 µg/mL), and Amphotericin B (0.25 µg/mL)

16. The maintenance protocol for the 22Rv1/MMTV_GR-KO cell line is based on the ATCC 22Rv1 maintenance protocol (11). 22Rv1/MMTV_GR-KO cells are maintained in a culture medium that includes 200 µg/mL hygromycin as a luciferase gene selection marker to be used the first time after thawing cells. 0.1%Trypsin-EDTA is preferred over 0.05% Trypsin-EDTA for passage of 22Rv1/MMTV_GR-KO cell line, because the higher concentration improves cell dissociation from the cell culture plate. For the assay, cells should be suspended at 3.0×10^5 cells per 1 mL with test medium. 100 µL aliquots of suspended cells (corresponding to 3.0×10^4 cells /well) should be transferred into a 96-well white plate. Cells are pre-incubated for 48 hours at 37 °C in a 5%±0.5% CO_2 incubator prior to exposure.

17. DCC-FBS in test medium is used to minimize the interference of other serum ingredients.

Vehicle control, AR agonist control and AR antagonist control

18. For the AR agonist assay, the agonist control (PC_{AGO1}) wells (n = 4) treated with a 10 nM DHT and vehicle control (VC) wells (n =4) containing only 0.1% DMSO, and cytotoxicity control (PC_{CT}; 1 mM SDS) wells (n = 4) should be prepared on each plate. The 10 nM DHT concentration is selected in order to achieve 100% response in the AR agonist assay.

19. For the AR antagonist assay, VC wells (n = 3), agonist control (PC_{AGO2}; 800 pM DHT) wells (n = 3), AR antagonist control (PC_{ANTA}; 800 pM DHT and 1 μM of Bicalutamide) wells (n = 3), and cytotoxicity control (PC_{CT}; 800 pM DHT and 1 mM SDS) wells (n = 3) should be included for each plate.

Positive and negative references standards

20. Reference standards for each assay should be included in one plate of each run. For the AR agonist assay, three well-characterised reference standards; two positive reference standards (DHT and Mestanolone) and one negative reference standard (Diethylhexyl phthalate (DEHP)) should be included. Reference standards for the AR antagonist assay include two positive reference standards (Bicalutamide and Bisphenol A) and one negative reference standard (DEHP).

Quality criteria for AR agonist/antagonist assay

21. The mean luciferase activity of the PC (AR agonist assay: 10 nM DHT (PC_{AGO1}); AR antagonist assay: 800 pM DHT (PC_{AGO2})) should be at least 13-fold greater than the mean VC on each plate for the AR agonist assay, and at least 10-fold greater than the mean VC for the AR antagonist assay. With respect to the quality control of the assay, the induction fold of the PC_{10} must be greater than 1 + 2 Standard Deviations (SD) of the induction of the VC. Relative transcriptional activity (RTA) of PC_{ANTA} (800 pM DHT and 1 μM Bicalutamide), which is a single concentration without a dose response curve of Bicalutamide, should be less than 53.6% of the PC_{AGO2} in the AR antagonist assay.

Acceptability criteria

Table E.1. Acceptability criteria for AR agonist assay

Chemicals	Log PC_{10}	Log PC_{50}	Test Range
5α-Dihydrotestosterone (DHT)	−12.2 to −9.7	−10.6 to −9.0	1.0×10^{-6} to 1.0×10^{-12} M
Mestanolone	−12.3 to −9.8	−10.2 to −8.6	1.0×10^{-6} to 1.0×10^{-12} M
Diethylhexyl phthalate (DEHP)	–	–	1.0×10^{-5} to 1.0×10^{-11} M

Induction fold of PC_{AGO1}	≥ 13
Induction fold of PC_{10}	Greater than 1+2SD (induction of VC)

Induction fold of PC_{10}: corresponding to the PC_{10} (10%) of AR agonist control (PC_{AGO1}:10 nM of DHT)
SD: Standard Deviation, VC: Vehicle Control

22. Induction fold of PC_{AGO1} is calculated using the following equation:

$$\text{Induction fold of } PC_{AGO1} = \frac{\text{Mean RLU of } PC_{AGO1} \text{ (10 nM DHT)}}{\text{Mean RLU of Vehicle control}}$$

▪ RLU: relative light units

Table E.2. Acceptability criteria for AR antagonist assay

Chemicals	Log IC_{30}	Log IC_{50}	Test Range
Bicalutamide	−7.5 to −6.2	−7.0 to −5.8	1.0×10^{-4} to 1.0×10^{-10} M
Bisphenol A	−6.6 to −5.4	−6.2 to −5.0	1.0×10^{-5} to 1.0×10^{-11} M
DEHP	-	-	1.0×10^{-5} to 1.0×10^{-11} M
Induction fold of PC_{AGO2}	≥ 10		
RTA of PC_{ANTA} (%)	≤53.6		

23. Induction fold of PC_{AGO2} is calculated using the following equation:

$$\text{Induction fold of } PC_{AGO2} = \frac{\text{Mean RLU of } PC_{AGO2} \text{ (800 pM DHT)}}{\text{Mean RLU of Vehicle control}}$$

$$\text{RTA of } PC_{ANTA} (\%) = \frac{\text{Mean RLU of } PC_{ANTA} - \text{Mean RLU of VC}}{\text{Mean RLU of } PC_{AGO2} - \text{Mean RLU of VC}} \times 100$$

▪ RTA: relative transcriptional activity

Solubility test

24. The solubility test is based on the OECD GIVIMP (13). Test chemical stocks are prepared at a maximum concentration of up to 1 M (stock solution; 0.1% of the stock

solution in wells with cells, i.e. 1 mM) in DMSO or an appropriate solvent. If precipitation occurs, the stock solution should be re-prepared a new concentration solution at 10 times lower than the original stock solution until no precipitation is observed.

Test chemical exposure and assay plate organisation

Pre-screen run in AR agonist assay

25. The maximal stock concentration of each test chemical, determined by the solubility test (see above), should be serially diluted at a ratio of 1:10 in DMSO (or another appropriate solvent). Then the dilutions are added to aqueous medium to achieve a final DMSO concentration of 0.1%. The recommended final volume for each well is 100 µL (the test medium from the assay plate should be removed and replaced with the test chemicals in the test medium). Triplicate wells are used for each concentration. The reference standards for the AR agonist assay (DHT, Mestanolone and DEHP) should be tested in every assay. Wells treated with 10 nM DHT (PC_{AGO1}), wells treated with 0.1% DMSO alone (VC) and wells treated with 1 mM SDS (PC_{CT}) should be included in each plate for the AR agonist assay (Table E.3). An example of the plate design of test chemicals is provided in Table E.4. After adding the test chemicals, the assay plates should be placed at 37°C±1°C in a 5%±0.5% CO_2 incubator for 20-24 hours.

Table E.3. Example of plate concentration assignment for the reference chemicals (in M).

	DHT			**Mestanolone**			**DEHP**			**Test Chemical**		
	1	2	3	4	5	6	7	8	9	10	11	12
A	$1.0x10^{-6}$	→	→	$1.0x10^{-6}$	→	→	$1.0x10^{-5}$	→	→	$1.0x10^{-3}$	→	→
B	$1.0x10^{-7}$	→	→	$1.0x10^{-7}$	→	→	$1.0x10^{-6}$	→	→	$1.0x10^{-4}$	→	→
C	$1.0x10^{-8}$	→	→	$1.0x10^{-8}$	→	→	$1.0x10^{-7}$	→	→	$1.0x10^{-5}$	→	→
D	$1.0x10^{-9}$	→	→	$1.0x10^{-9}$	→	→	$1.0x10^{-8}$	→	→	$1.0x10^{-6}$	→	→
E	$1.0x10^{-10}$	→	→	$1.0x10^{-10}$	→	→	$1.0x10^{-9}$	→	→	$1.0x10^{-7}$	→	→
F	$1.0x10^{-11}$	→	→	$1.0x10^{-11}$	→	→	$1.0x10^{-10}$	→	→	$1.0x10^{-8}$	→	→
G	$1.0x10^{-12}$	→	→	$1.0x10^{-12}$	→	→	$1.0x10^{-11}$	→	→	$1.0x10^{-9}$	→	→
H	**VC**	→	→	→	**PC_{AGO1}**	→	→	→	**PC_{CT}**	→	→	→

- VC: Vehicle control (0.1% DMSO)
- PC_{AGO1}: AR agonist control (10 nM DHT)
- PC_{CT}: Cytotoxic control (1 mM SDS)

Table E.4. Example of plate concentration assignment for the test chemicals (in M).

	Test Chemical 1			**Test Chemical 2**			**Test Chemical 3**			**Test Chemical 4**		
	1	2	3	4	5	6	7	8	9	10	11	12
A	$1.0x10^{-3}$	→	→	$1.0x10^{-3}$	→	→	$1.0x10^{-5}$	→	→	$1.0x10^{-6}$	→	→
B	$1.0x10^{-4}$	→	→	$1.0x10^{-4}$	→	→	$1.0x10^{-6}$	→	→	$1.0x10^{-7}$	→	→
C	$1.0x10^{-5}$	→	→	$1.0x10^{-5}$	→	→	$1.0x10^{-7}$	→	→	$1.0x10^{-8}$	→	→
D	$1.0x10^{-6}$	→	→	$1.0x10^{-6}$	→	→	$1.0x10^{-8}$	→	→	$1.0x10^{-9}$	→	→
E	$1.0x10^{-7}$	→	→	$1.0x10^{-7}$	→	→	$1.0x10^{-9}$	→	→	$1.0x10^{-10}$	→	→
F	$1.0x10^{-8}$	→	→	$1.0x10^{-8}$	→	→	$1.0x10^{-10}$	→	→	$1.0x10^{-11}$	→	→
G	$1.0x10^{-9}$	→	→	$1.0x10^{-9}$	→	→	$1.0x10^{-11}$	→	→	$1.0x10^{-12}$	→	→
H	**VC**	→	→	→	**PC_{AGO1}**	→	→	→	**PC_{CT}**	→	→	→

- VC: Vehicle control (0.1% DMSO)
- PC_{AGO1}: AR agonist control (10 nM DHT)
- PC_{CT}: Cytotoxic control (1 mM SDS)

Comprehensive run in AR agonist assay

26. The test chemicals, which are determined to be an AR agonist in the pre-screen run should be further tested with a comprehensive run. The maximal concentration of the test chemical, determined from the concentration response curve generated in the pre-screen run, should be serially diluted at a ratio of 1:3 or 1:5 in DMSO (see Appendix E.1). These dilutions are then added to aqueous medium to a final DMSO concentration of 0.1%, and all concentrations should be tested in triplicate. All tests should be conducted at concentrations where the concentration–response curve can be well characterised. To achieve these conditions, solutions found to contain insoluble solids or concentrations found to induce cytotoxic effects against cell lines should not be included in the final analysis. The recommended final volume for each well is 100 µL (test medium from assay plate should be removed and replaced with test chemicals in test medium). The plate layout for the reference standards and the test chemicals run in the comprehensive run is the same as for the pre-screen run. After adding the test chemicals, the assay plates should be placed at 37°C±1°C in a 5%±0.5% CO_2 incubator for 20-24 hours.

Pre-screen run in AR antagonist assay

27. The maximal stock concentration of each test chemical, determined by the solubility test (see above), should be serially diluted at a ratio of 1:10 in DMSO. These dilutions are then added to aqueous medium to a final DMSO concentration of 0.1%. The recommended final volume for each well is 100 µL (test medium from the assay plate should be removed and replaced with the test chemicals in the test medium). The AR antagonist assay reference standards (Bicalutamide, Bisphenol A and DEHP) should be tested in every assay. An AR agonist control (PC_{AGO2}; 800 pM DHT), an AR antagonist control (PC_{ANTA}; 800 pM DHT and 1 µM Bicalutamide) and cytotoxic control (PC_{CT}; 800

pM DHT and 1 mM SDS) should be prepared for the AR antagonist assay (Table E.5). The plate design of the test chemicals is provided in Table E.6. Except for the VC, all other wells are spiked with a fixed concentration of the agonistic reference chemical (800 pM DHT) in order to measure attenuation of the agonistic response. After adding the test chemicals, the assay plates should be placed at 37°C±1°C in a 5%±0.5% CO_2 incubator for 20-24 hours.

Table E.5. Example of plate concentration assignment for the reference standards (in M)

	Bicalutamide			**Bisphenol A**			**DEHP**			**Test Chemical 1**		
	1	2	3	4	5	6	7	8	9	10	11	12
A	$1.0x10^{-4}$	→	→	$1.0x10^{-5}$	→	→	$1.0x10^{-5}$	→	→	$1.0x10^{-3}$	→	→
B	$1.0x10^{-5}$	→	→	$1.0x10^{-6}$	→	→	$1.0x10^{-6}$	→	→	$1.0x10^{-4}$	→	→
C	$1.0x10^{-6}$	→	→	$1.0x10^{-7}$	→	→	$1.0x10^{-7}$	→	→	$1.0x10^{-5}$	→	→
D	$1.0x10^{-7}$	→	→	$1.0x10^{-8}$	→	→	$1.0x10^{-8}$	→	→	$1.0x10^{-6}$	→	→
E	$1.0x10^{-8}$	→	→	$1.0x10^{-9}$	→	→	$1.0x10^{-9}$	→	→	$1.0x10^{-7}$	→	→
F	$1.0x10^{-9}$	→	→	$1.0x10^{-10}$	→	→	$1.0x10^{-10}$	→	→	$1.0x10^{-8}$	→	→
G	$1.0x10^{-10}$	→	→	$1.0x10^{-11}$	→	→	$1.0x10^{-11}$	→	→	$1.0x10^{-9}$	→	→
H	**VC**			**PC_{AGO2}**			**PC_{ANTA}**			**PC_{CT}**		

- VC: Vehicle control (0.1% DMSO)
- PC_{AGO2}: AR Agonist control for AR antagonist assay (800 pM DHT)
- PC_{ANTA}: AR Antagonist control (1 µM Bicalutamide)
- PC_{CT}: Cytotoxic control (1 mM SDS)
- Grey wells include 800 pM DHT

Table E.6. Example of plate concentration assignment for test chemicals (in M)

	Test Chemical 1			Test Chemical 2			Test Chemical 3			Test Chemical 4		
	1	2	3	4	5	6	7	8	9	10	11	12
A	$1.0x10^{-3}$	→	→	$1.0x10^{-3}$	→	→	$1.0x10^{-5}$	→	→	$1.0x10^{-3}$	→	→
B	$1.0x10^{-4}$	→	→	$1.0x10^{-4}$	→	→	$1.0x10^{-6}$	→	→	$1.0x10^{-4}$	→	→
C	$1.0x10^{-5}$	→	→	$1.0x10^{-5}$	→	→	$1.0x10^{-7}$	→	→	$1.0x10^{-5}$	→	→
D	$1.0x10^{-6}$	→	→	$1.0x10^{-6}$	→	→	$1.0x10^{-8}$	→	→	$1.0x10^{-6}$	→	→
E	$1.0x10^{-7}$	→	→	$1.0x10^{-7}$	→	→	$1.0x10^{-9}$	→	→	$1.0x10^{-7}$	→	→
F	$1.0x10^{-8}$	→	→	$1.0x10^{-8}$	→	→	$1.0x10^{-10}$	→	→	$1.0x10^{-8}$	→	→
G	$1.0x10^{-9}$	→	→	$1.0x10^{-9}$	→	→	$1.0x10^{-11}$	→	→	$1.0x10^{-9}$	→	→
H	**VC**			**PC_{AGO2}**			**PC_{ANTA}**			**PC_{CT}**		

- VC: Vehicle control (0.1% DMSO)
- PC_{AGO2}: AR Agonist control for AR antagonist assay (800 pM DHT)
- PC_{ANTA}: AR Antagonist control (1 µM Bicalutamide)
- PC_{CT}: Cytotoxic control (1 mM SDS)
- Grey wells include 800 pM DHT

Comprehensive run and specificity control test in AR antagonist assay

28. To ensure the identification of AR antagonist that is determined to be positive in the pre-screen run, the comprehensive run and specificity control test should be conducted using both 800 pM DHT and 100 nM DHT. The inclusion of these two concentrations of DHT in the antagonist assay is expected to result in a shift between the concentration-response curves of "true" AR antagonists and distinguish these chemicals from potential false positives. The maximal concentration of the test chemical, determined from the concentration–response curves generated in the pre-screen run, should be serially diluted at a ratio of 1:3 or 1:5 in DMSO (see Appendix E.1). These dilutions are then added to aqueous medium to a final DMSO concentration of 0.1%, and all concentrations should be tested in triplicate. The recommended final volume for each well is 100 µL(the test medium from the assay plate should be removed and replaced with the test chemicals in test medium). The plate layout for the reference standards is the same as for the pre-screen run and the plate layout for the test chemicals is shown in Table E.7. An AR agonist control (PC_{AGO2}; 800 pM DHT), an AR antagonist control (PC_{ANTA}; 800 pM DHT and 1 µM Bicalutamide) and cytotoxic control (PC_{CT}; 800 pM DHT and1 mM SDS) should be prepared for the AR antagonist assay. The plate layout is given in Table E.7. After adding the test chemicals, the assay plates should be placed at 37°C±1°C in a 5%±0.5% CO_2 incubator for 20-24 hours.

Table E.7. Example of plate concentration assignment of test chemicals (in log M)

	Test chemical 1						Test chemical 2					
	1	2	3	4	5	6	7	8	9	10	11	12
A	−5	→	→	−5	→	→	−4	→	→	−4	→	→
B	−5.7	→	→	−5.7	→	→	−4.7	→	→	−4.7	→	→
C	−6.4	→	→	−6.4	→	→	−5.4	→	→	−5.4	→	→
D	−7.1	→	→	−7.1	→	→	−6.1	→	→	−6.1	→	→
E	−7.8	→	→	−7.8	→	→	−6.8	→	→	−6.8	→	→
F	−8.5	→	→	−8.5	→	→	−7.5	→	→	−7.5	→	→
G	−9.2	→	→	−9.2	→	→	−8.2	→	→	−8.2	→	→
H	**VC**			**PC_{AGO2}**			**PC_{ANTA}**			**PC_{CT}**		

- VC: Vehicle control (DMSO);
- PC_{AGO2}: AR agonist control (800 pM of DHT);
- PC_{ANTA}: AR antagonist control (1 μM of Bicalutamide);
- PC_{CT}: Cytotoxicity control (1 mM of SDS);
- Grey wells are spiked with 800 pM DHT;
- Dark grey wells are spiked with 100 nM DHT

Endpoint measurements

29. Endpoint are measured using the Steady-Glo Luciferase assay system (e.g. Promega, E2510, or equivalents) for AR response, and the live-cell protease detection system (e.g. Cell Titer-Fluor™ Cell viability assay, Promega, G6080, or equivalents) for the cytotoxicity. The measurements of cell viability and luciferase activity are performed in the same plate.

30. For cell viability assay:

- Prepare the cell viability (CellTiter-Fluor™) reagent according to the manufacturer's instructions.
- Add directly 20 μL/well of cell viability assay reagent into the assay wells containing medium with test chemicals.
- Mix the assay plates briefly using an orbital shaker.
- Incubate the assay plates at 37°C±1°C in a 5%±0.5% CO_2 incubator for 1–3 hour.
- Remove plates from incubator and measure the cytotoxicity using a fluorometer (380–400 nm Ex /505 nm Em).

31. For luciferase assay

- Prepare the luciferase assay (Steady-Glo) reagent according to the manufacturer's instructions.
- Add directly 50 μL/well of luciferase assay reagent into the assay wells after the cell viability

assay.

- Cover the top of the assay plate with aluminium foil to block the light, and leave at room temperature for 5-10 min.
- Measure the luciferase activity using a luminescence reader.

Analysis of Data

Cytotoxicity

32. Cytotoxicity, as read by the fluorometer in RFU units, is recorded and is transformed as follows:

- The average for the AR agonist and AR antagonist control (AR agonist assay: 10 nM DHT, AR antagonist assay: 800 pM DHT) is set at 100%.
- The average for cytotoxicity control (AR agonist assay: 1 mM SDS, AR antagonist assay: 80 0 pM DHT and 1 mM SDS) is set at 0%

33. If the results of the cell viability test indicate that the concentration of the test chemical has reduced cell viability by 20% or more, this concentration is regarded as cytotoxic. All concentrations considered cytotoxic should be excluded from the evaluation

34. For the cell viability assay, the data transformation from RFU units is as follows:

$$\text{Cell viability (\%)} = \frac{\text{Mean RFU of test chemical - Mean RFU of } PC_{CT}}{\text{Mean RFU of PC – Mean RFU of } PC_{CT}} \times 100$$

▪ RFU: relative fluorescence units

Luciferase activity

35. The luminescence signal data, as read by the luminometer in RLU units, is recorded and is transformed as follow:

- The average for the AR agonist and AR antagonist control (AR agonist assay: 10 nM DHT, AR antagonist assay: 800 pM DHT) is set at 100%.
- The average for vehicle control (0.1% DMSO) is set at 0%

36. For the agonist, and antagonist assay, the data transformation from RLU units is as follows:

$$\text{RTA (\%)} = \frac{\text{Mean RLU of test chemical - Mean RLU of VC}}{\text{Mean RLU of PC – Mean RLU of VC}} \times 100$$

▪ RLU: relative light units▪ RTA: relative transcriptional activity

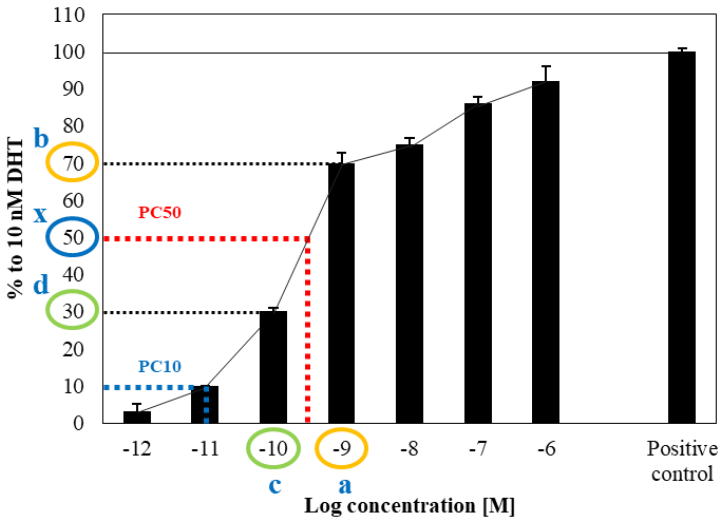
Calculation of parameters

37. In the AR agonist assay, the following information should be provided for a positive test chemical: the concentrations that induce an effect corresponding to that of a 10% effect for the positive control (log PC_{10}) and, if appropriate, the 50% effect for the positive control (log PC_{50}). Descriptions of log PC_x values, where "x" is a selected response, e.g. 10% or 50% induction, compared to PC_{AGO1}, are provided in Figure E.2. Log PC_{10} and log PC_{50} values can be defined as the test chemical concentrations estimated to

elicit either a 10% or a 50% induction of transcriptional activity by PC_{AGO1} (10 nM of DHT). Each log PC_x value can be calculated by a simple linear regression using two variable data points for the transcriptional activity. Where the data points lying immediately above and below the log PC_x value have the coordinates (a, b) and (c, d) respectively, then the log PC_x value is calculated using the equation below and Figure E. 2 shows the method for the calculation of log[PC_{50}]:

$$\log[PC_x] = c+[(x-d)/(b-d)](a-c)$$

Figure E.2. Schematic illustration of the calculation of log PC_x values

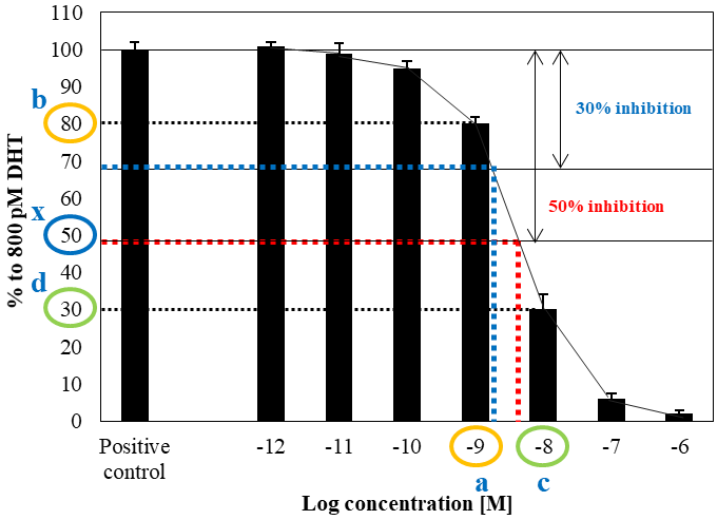


▪ The PC_{AGO1} (10 nM of DHT) is included on each assay plate in AR agonist assay.

38. For the AR antagonist assay, the following information should be provided for a positive test chemical: the concentrations for 30% inhibition of transcriptional activity induced by 800 pM DHT (log IC_{30}) and, if appropriate, for 50% inhibition of activity by 800 pM DHT (log IC_{50}). Descriptions of log IC_x values, where "x" is a selected response, e.g. 30% or 50% inhibition, compared to PC_{AGO2}, are provided in Figure E.3. Log IC_{50} and log IC_{30} values can be defined as the test chemical concentrations estimated to elicit either a 50% or a 30% inhibition of transcriptional activity induced by 800 pM DHT. Each log IC_x value can be calculated by a simple linear regression using two variable data points for the transcriptional activity. Where the data points lying immediately above and below the log IC_x value have the coordinates (c, d) and (a, b) respectively, then the log IC_x value is calculated using the equation below and Figure E.3 shows an illustration of the calculation of log[IC_{50}]:

$$\log [IC_x] = a-[(b-(100-x))/(b-d)](a-c)$$

Figure E.3. Schematic illustration of the calculation of log IC_x values



▪ The PC_{AGO2} (800 pM DHT) is included on each assay plate in AR antagonist assay.

39. In case of the specificity control test, to distinguish the responses by the two concentrations of DHT, the Y_C represents the relative induction at concentration c when the 800 pM DHT is used, and the symbol S_C represents the relative induction at concentration c when the 100 nM DHT is used. The data transformation from RLU of Y_C or S_C is as follows:

$$Y_C \text{ or } S_C\ (\%) = \frac{\text{Mean RLU of test chemical} - \text{Mean RLU of VC}}{\text{Mean RLU of } PC_{AGO2} - \text{Mean RLU of VC}} \times 100$$

40. For test chemicals to be a true AR antagonist (competitive), the square of the coefficient of determination, R^2, was calculated between the relative induction of the standard response Y_C and the relative induction of the specificity response S_C. If R^2 is less than 0.9, this test chemical was determined to be a true AR antagonist. The formula of R^2 for identifying the true AR antagonist can be found in the validation report (6). Some caution should be applied as this criterion cannot be considered as 100% definitive (as shown in the AR-CALUX® validation study report). It may be influenced by the shape of the curves and by outliers. Expert judgment may need to be applied.

41. The presence of increasing levels of cytotoxicity can significantly alter or eliminate the typical sigmoidal response and should be considered when interpreting the data in the agonist and antagonist assay. Accordingly, AR-mediated transcriptional activity and cytotoxicity should be evaluated simultaneously in the same assay plate. Should the results of the cytotoxicity test show that the concentration of the test chemical has reduced cell viability by 20% or more, this concentration is regarded as cytotoxic, and the concentrations at or above the cytotoxic concentration should be excluded from the evaluation.

Data Interpretation Criteria

42. The interpretation of data and the decision, whether a test chemical in the absence of cytotoxicity is considered positive or negative, are shown in Table E.8.

43. To classify a chemical as an AR agonist, a positive pre-screen run in which a log PC_{10} can be determined should be followed by concordant results in two comprehensive runs or if not concordant, a third comprehensive run. In the case of a negative pre-screen run, the result should be confirmed in a (second) follow-up pre-screen run. If the second pre-screen run is positive after a first negative pre-screen run, a third pre-screen run should be additionally conducted. In the case of AR antagonist, the log IC_{30} is calculated in a pre-screen run and is confirmed in at least two (of up to three) comprehensive runs (in the absence of cytotoxicity) alongside a specificity control test. If the R^2 of test chemical in the specificity control is less than 0.9, the test chemical can be considered a true AR antagonist, however this may require additional expert judgement (see paragraph 40). Chemicals that are not AR antagonists are classified based on negative results (in the absences of positive results) in at least two pre-screen runs (Table E.8).

Table E.8. Positive and negative decision criteria

AR agonist assay	**Positive**	If obtained RPC_{max} is equal to or exceeds 10% of the response of the positive control.
	Negative	In all other cases.
AR antagonist assay	**Positive**	If the test chemical satisfies the following: i) the log IC_{30} of test chemical is calculated in the absence of cytotoxicity and ii) the R^2 is less than 0.9 in specificity control test.
	Negative	In all other cases.

▪ All results are in the absence of cytotoxicity.

Literature

1. Kim H.J., Park Y.I., Dong M.S. (2006) Comparison of prostate cancer cell lines for androgen receptor-mediated reporter gene assays. *Toxicology In Vitro* 20, 1159–1167.

2. Sun S, Park E.J, Choi Y.H., Lee H.S., Ahn B.Y. and Dong M.S. 2016. Development and prevalidation of an in vitro transactivation assay for detection of (anti)androgenic potential compounds using 22Rv1/ MMTV cells. *Reproductive Toxicology* 60, 156–166.

3. Yen P.M., Liu Y., Palvimo J.J., Trifiro M., Whang J., Pinsky L., Janne O.A. and Chin W.W. 1997. Mutant and wild-type androgen receptors exhibit cross-talk on androgen- glucocorticoid-, and progesterone-mediated transcription. *Molecularl Endocrinology* 11, 162-171.

4. Zwart, N., Andringa, D., de Leeuw, W.J., Kojima, H., Iida, M., Houtman, C. 2017. Improved androgen specificity of AR-Ecoscreen by CRISPR based glucocorticoid receptor knockout. *Toxicology in Vitro* 45, 1-9.

5. Lee H.S., Lee S.H. and Park Y.H. 2019. Enhancement of androgen transcriptional activation assay based on genome edited glucocorticoid knock out human prostate cancer cell line. *Environmental Research* 171, 437-443.

6. Validation Study Report of 22Rv1/MMTV_GR-KO ARTA assay. (2019). Available at (http://www.nifds.go.kr/brd/m_18/view.do?seq=12486&srchFr=&srchTo=&srchWord=&srchTp=&itm_seq_1=0&itm_seq_2=0&multi_itm_seq=0&company_cd=&company_nm=&page=1)

7. OECD (2012) Guidance document on standardised test guidelines for evaluating chemicals for endocrine disruption. OECD Environment, Health and Safety Publications Series on Testing and Assessment. No. 150. ENV/JM/MONO(2012)22.

8. OECD (2005) Guidance document on the validation and international acceptance of new or updated test methods for hazard assessment. OECD Environment, Health and Safety Publications Series on Testing and Assessment. No. 34. ENV/JM/MONO(2005)14.

9. OECD (2016) OECD Guideline for the Testing of Chemicals: Stably Transfected Human Androgen Receptor Transcriptional Activation Assay for Detection of Androgenic Agonist and Antagonistic Activity of Chemicals.

10. Validation report of the 22Rv1/MMTV_GR-KO AR TA assay: An *in vitro* assay for identifying human androgen receptor agonists and antagonists.

11. ATCC 22Rv1 culture method. Available at (https://www.atcc.org/products/all/CRL-2505.aspx#culture method)

12. Kobayashi, H., Yamamoto, K., Eguchi, M., Kubo, M., Nakagami, S., Wakisaka, S., Kaizuka, M. and Ishii H (1995), Rapid detection of mycoplasma contamination in cell cultures by enzymatic detection of polymerase chain reaction (PCR) products. *J. Vet. Med. Sci.* 57, 769–771.

13. OECD (2018) Guidance document on good in vitro method practices (GIVIMP). OECD series on Testing and Assessment. No. 286. OECD Publishing Paris.

Appendix E.1

1. The method to determine the max concentration for the comprehensive run

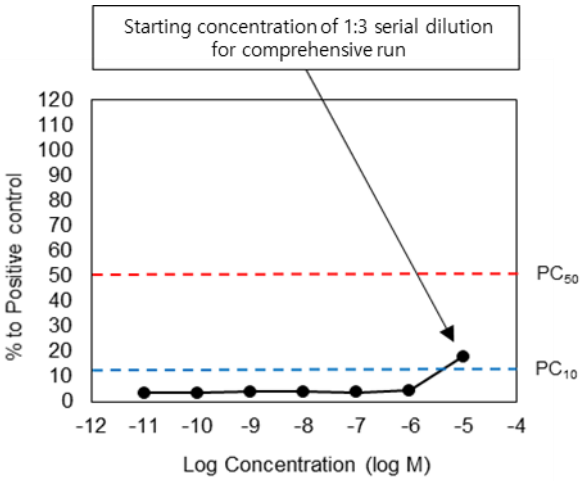
2. If test chemicals are determined to be positive in the pre-screen run, comprehensive run should be conducted to accurately determine the potency of test chemicals. All test chemicals classified as positive for AR agonistic activity should have a concentration–response curve consisting of a baseline, and a positive slope; all test chemicals classified as positive for AR antagonistic activity should have a concentration response curve consisting of a baseline, and a negative slope. If possible, PC_{10}, PC_{50}, IC_{30} and IC_{50} value should be calculated for each positive decision. The comprehensive AR agonist/antagonist assay consists of a seven-point serial dilution (1:3 or 1:5 serial dilution) with each concentration tested in triplicate wells of the 96-well plate. To determine the starting concentrations for comprehensive run, use the following criteria:

- If results in the pre-screen run suggest that the test chemical is positive with only PC_{10} value for AR agonist assay (if there is only one point on the test chemicals concentration curve that is greater than the positive decision criteria without cytotoxicity), the comprehensive run should be conducted using the 7-point 1:3 serial dilution starting at the maximum exposure concentration (see example 1).

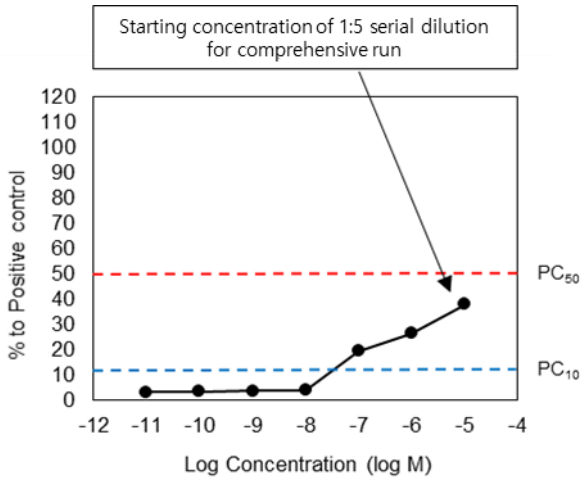
- If results in the pre-screen run suggest that the test chemical is positive with only PC_{10} value for AR agonist assay (if there are several points on the test chemical concentration curve that are greater than the positive decision criteria without cytotoxicity), the comprehensive run should be conducted using the 7-point 1:5 serial dilution starting at the maximum exposure concentration (see example 2).

- If results in the pre-screen run suggest that the test chemical is positive with PC_{10} and PC_{50} values for AR agonist assay (i.e., if there are points on the test chemical concentration curve that are greater than the positive decision criteria without cytotoxicity), the starting concentration to be used for the 7-point dilution scheme in the comprehensive run should be 10 times greater than the concentration associated with the highest level of response in the pre-screen run (see example 3).

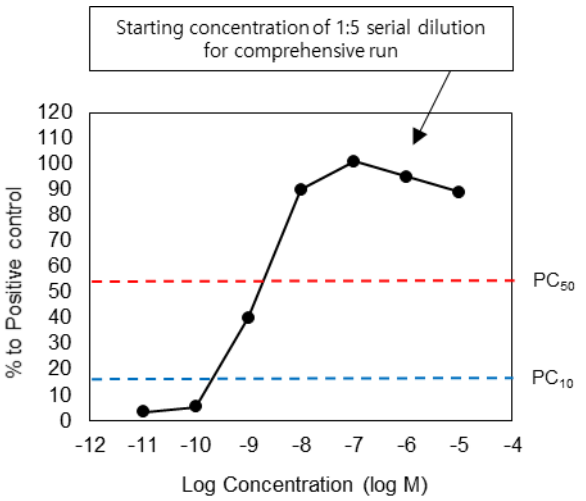
- If results in the pre-screen run suggest that the test chemical is positive with only PC_{50} value (or a PC_{50} value cannot be calculated but maximum activity is more than 10%) for AR agonist assay (i.e., if all testing points on the test chemical concentration curve are greater than the positive decision criteria without cytotoxicity), the starting concentration to be used for the 7-point dilution scheme in the comprehensive run should be 10 times greater than concentration associated with the highest level of response in the pre-screen run (see example 4).



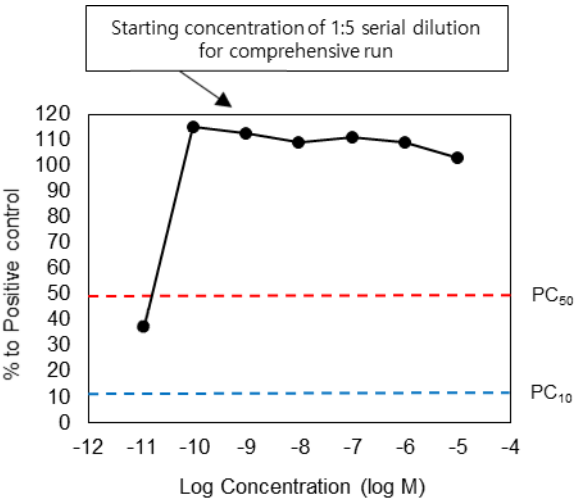
Example 1



Example 2



Example 3



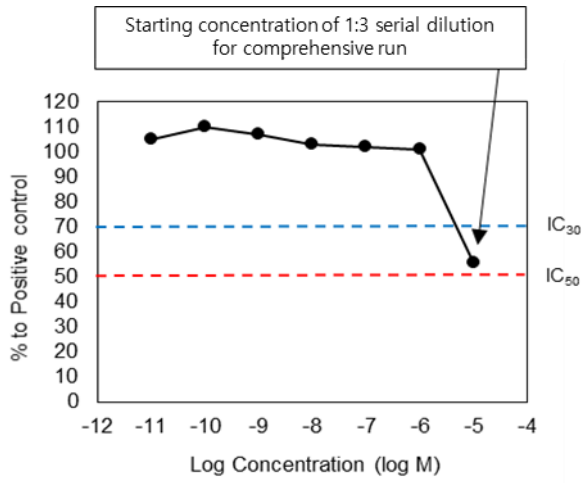
Example 4

- If results in the pre-screen run suggest that the test chemical is positive with only IC_{30} value for AR antagonist assay (if there is only one point on the test chemical concentration curve for which the response is greater than the positive decision criteria without cytotoxicity), the comprehensive testing will be conducted using the 7-point 1:3 serial dilution starting at the maximum exposure concentration (see example 5).
- If results in the pre-screen run suggest that the test chemical is positive with only IC_{30} value for AR antagonist assay (if there are points on the test chemical concentration curve for which the response is greater than the positive decision criteria without cytotoxicity), the comprehensive testing should be conducted using the 7-point 1:5 serial dilution starting at the maximum exposure concentration (see example 6).

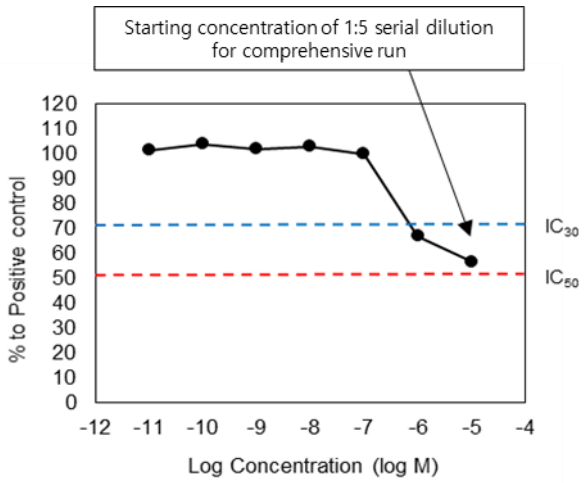
- If results in the pre-screen run suggest that the test chemical is positive with IC_{30} and IC_{50} values for AR antagonist assay (i.e., if there are points on the test chemical concentration curve that have a response greater than the positive decision criteria without cytotoxicity), the starting concentration to be used for the 7-point dilution scheme in the comprehensive testing should be the concentration giving the highest level of response in the pre-screen run (see example 7).

- If results in the pre-screen run suggest that the test chemical is positive with only IC_{50} value (or not calculate IC_{50} value) for AR antagonist assay (i.e., if all testing points on the test chemical concentration curve are greater than the positive decision criteria without cytotoxicity), the starting concentration to be used for the 7-point dilution scheme in the comprehensive testing should be the concentration giving the highest level of response in the pre-screen run (see example 8).

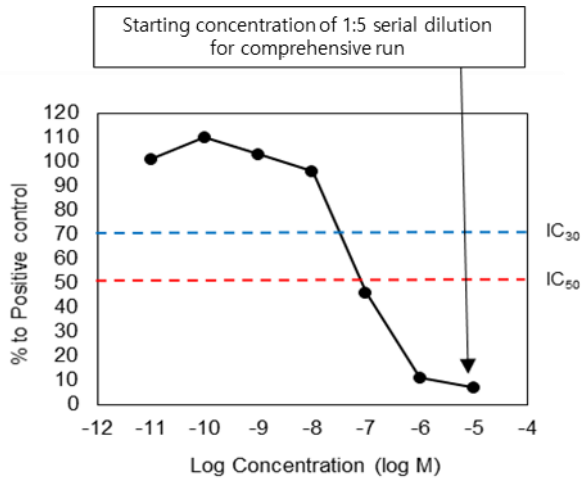
-



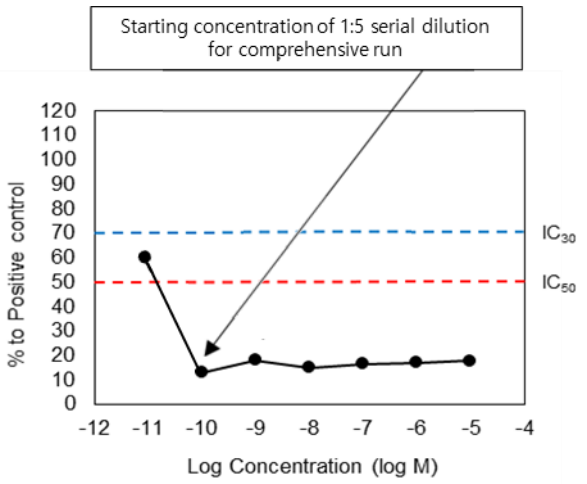
Example 5



Example 6



Example 7



Example 8

Section 4
Health effects

Test Guideline No. 491

Short Time Exposure *In Vitro* Test Method for Identifying i) Chemicals Inducing Serious Eye Damage and ii) Chemicals Not Requiring Classification for Eye Irritation or Serious Eye Damage

26 June 2020

OECD Guidelines for the Testing of Chemicals

Adopted:
26 June 2020

OECD GUIDELINE FOR TESTING OF CHEMICALS

DRAFT UPDATED TEST GUIDELINE 491: SHORT TIME EXPOSURE IN VITRO TEST METHOD FOR EYE HAZARD POTENTIAL

INTRODUCTION

1. The Short Time Exposure (STE) test method is an in vitro method that can be used under certain circumstances and with specific limitations for hazard classification and labeling of chemicals (substances and mixtures) that induce serious eye damage as well as those that do not require classification for either serious eye damage or eye irritation, as defined by the United Nations (UN) Globally Harmonized System of Classification and Labeling of Chemicals (GHS) (1).

2. For many years, the eye hazard potential of chemicals has been evaluated primarily using an in vivo rabbit eye test (TG 405). It is generally accepted that, in the foreseeable future, no single in vitro alternative test will be able to fully replace the in vivo rabbit eye test to predict across the full range of serious eye damage/eye irritation responses for different chemical classes. However, strategic combinations of alternative test methods used in a (tiered) testing strategy may well be able to fully replace the rabbit eye test (2). The top-down approach is designed for the testing of chemicals that can be expected, based on existing information, to have a high irritancy potential or induce serious eye damage. Conversely, the bottom-up approach is designed for the testing of chemicals that can be expected, based on existing information, not to cause sufficient eye irritation to require a classification. While the STE test method is not considered to be a complete replacement for the in vivo rabbit eye test, it is suitable for use as part of a tiered testing strategy for regulatory classification and labeling, such as the top-down/bottom-up approach, to identify without further testing (i) chemicals inducing serious eye damage (UN GHS Category 1) and (ii) chemicals (excluding all solid chemicals other than surfactants) that do not require classification for eye irritation or serious eye damage (UN GHS No Category) (1) (2). However, a chemical that is neither predicted to cause serious eye damage (UN GHS Category 1) nor UN GHS No Category (does not induce either serious eye damage or eye irritation) by the STE test method would require additional testing to establish a definitive classification. Furthermore, the appropriate regulatory authorities should be consulted before using the STE in a bottom-up approach under classification schemes other than the UN GHS. The choice of the most appropriate test method and the use of this Test Guideline should be seen in the context of the OECD Guidance Document on an Integrated Approaches on Testing and Assessment for Serious Eye Damage and Eye irritation (14).

3. The purpose of this test guideline (TG) is to describe the procedures used to evaluate the eye hazard potential of a test chemical based on its ability to induce cytotoxicity in the Short Time Exposure Test method. The cytotoxic effect of chemicals on corneal epithelial cells is an important mode of action (MOA) leading to corneal epithelium damage and eye irritation. Cell viability in the STE test method is assessed by the

quantitative measurement, after extraction from cells, of blue formazan salt produced by the living cells by enzymatic conversion of the vital dye MTT (3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide), also known as Thiazolyl Blue Tetrazolium Bromide (3). The obtained cell viability after 5 minutes exposure is compared to the solvent control (relative viability) and used to estimate the potential eye hazard of the test chemical. A test chemical is classified as UN GHS Category 1 when both the 5% and 0.05% concentrations result in a cell viability smaller than or equal to (≤) 70%. Conversely, a chemical is predicted as UN GHS No Category when both 5% and 0.05% concentrations result in a cell viability higher than (>) 70%.

4. The term "test chemical" is used in this Test Guideline to refer to what is tested and is not related to the applicability of the STE test method to the testing of substances and/or mixtures. Definitions are provided in Annex I.

INITIAL CONSIDERATIONS AND LIMITATIONS

5. This Test Guideline is based on a protocol developed by Kao Corporation (4), which was the subject of two different validation studies: one by the Validation Committee of the Japanese Society for Alternative to Animal Experiments (JSAAE) (5) and another by the Japanese Center for the Validation of Alternative Methods (JaCVAM) (6). A peer review was conducted by NICEATM/ICCVAM based on the validation study reports and background review documents on the test method (7).

6. When used to identify chemicals (substances and mixtures) inducing serious eye damage (UN GHS Category 1 (1), data obtained with the STE test method on 125 chemicals (including both substances and mixtures), showed an overall accuracy of 83% (104/125), a false positive rate of 1% (1/86), and a false negative rate of 51% (20/39) as compared to the in vivo rabbit eye test (7). The false negative rate obtained is not critical in the present context, since all test chemicals that induce a cell viability of ≤ 70% at a 5% concentration and > 70% at 0.05% concentration (see Table 2: Prediction model below) would be subsequently tested with other adequately validated in vitro test methods or, as a last option, in the in vivo rabbit eye test, depending on regulatory requirements, and in accordance with the sequential testing strategy and weight-of-evidence approaches currently recommended (1) (8). Mainly mono-constituent substances were tested, although a limited amount of data also exist on the testing of mixtures. The test method is nevertheless technically applicable to the testing of multi-constituent substances and mixtures. When considering testing of mixtures, difficult-to-test chemicals (e.g. unstable), or test chemicals not clearly within the applicability domain described in this Guideline, upfront consideration should be given to whether the results of such testing will yield results that are meaningful scientifically. The STE test method showed no other specific shortcomings when used to identify test chemicals as UN GHS Category 1. Investigators could consider using this test method on test chemicals, whereby cell viability ≤ 70% at both 5% and 0.05% concentration should be accepted as indicative of a response inducing serious eye damage that should be classified as UN GHS Category 1 without further testing.

7. When used to identify chemicals (substances and mixtures) not requiring classification for eye irritation and serious eye damage (i.e. UN GHS No Category), data obtained with the STE test method on 130 chemicals (including both substances and mixtures), showed an overall accuracy of 85% (110/130), a false negative rate of 12% (9/73), and a false positive rate of 19% (11/57) as compared to the in vivo rabbit eye test (7). If highly volatile substances (i.e. measured vapour pressure > 6kPa) and solid substances other than surfactants are excluded from the dataset, the overall accuracy

improves to 90% (92/102), the false negative rate to 2% (1/54), and the false positive to 19% (9/48) (7). Further work demonstrated that highly volatile substances can be correctly tested when using mineral oil instead of saline as a solvent (15). The accuracy of the STE test for highly volatile substances (i.e. vapour pressure > 6kPa) was then 95% (19/20), the false negative rate was 0% (0/7), and the false positive rate was 8% (1/13). As a consequence, the potential shortcoming of the STE test method when used to identify test chemicals not requiring classification for eye irritation and serious eye damage (UN GHS No Category) is a high false negative rate for solid chemicals (substances and mixtures) other than surfactants and mixtures composed only of surfactants. Such chemicals are excluded from the applicability domain of the STE test method (7). To that extent possible, test chemicals that are sensitive to hydrolysis should be evaluated under conditions that do not promote hydrolysis in order to avoid possible false negative results.

8. In addition to the chemicals mentioned in paragraphs 6 and 7, the STE test method generated dataset also contains in-house data on 40 mixtures, which when compared to the in vivo Draize eye test, showed an accuracy of 88% (35/40), a false positive rate of 50% (5/10), and a false negative rate of 0% (0/30) for predicting mixtures that do not require classification under the UN GHS classification system (9). The STE test method can therefore be applied to identify mixtures as UN GHS No Category in a bottom-up approach with the exception of solid mixtures other than those composed only of surfactants as an extension of its limitation to solid substances. Furthermore, mixtures containing substances with vapour pressure higher than 6kPa that do not dissolve in mineral oil, or that do not form stable suspensions for at least 5 minutes, are not currently within the applicability domain of the test method and may result in false negative outcomes.

9. The STE test method cannot be used for the identification of test chemicals as UN GHS Category 2, Category 2A (eye irritation) or UN GHS Category 2B (mild eye irritation), due to the considerable number of UN GHS Category 1 chemicals under-predicted as UN GHS Category 2, 2A, or 2B and UN GHS No Category chemicals over-predicted as UN GHS Category 2, 2A, or 2B (7). For this purpose, further testing with another suitable method may be required.

10. The STE test method is suitable for test chemicals that are dissolved or uniformly suspended for at least 5 minutes in physiological saline, 5% dimethyl sulfoxide (DMSO) in saline, or mineral oil (see paragraph 17 for solvent choice). The STE test method is not suitable for test chemicals that are insoluble or cannot be uniformly suspended for at least 5 minutes in physiological saline, 5% DMSO in saline, or mineral oil. The use of mineral oil in the STE test method is possible because of the short-time exposure. Therefore, the STE test method is suitable for predicting the eye hazard potential of water-insoluble test chemicals (e.g., long-chain fatty alcohols or ketones) provided that they are miscible in at least one of the three above proposed solvents (4).

PRINCIPLE OF THE TEST

11. The STE test method is a cytotoxicity-based in vitro assay that is performed on a confluent monolayer of Statens Seruminstitut Rabbit Cornea (SIRC) cells, cultured on a 96-well polycarbonate microplate (4). After five-minute exposure to both a 5% and a 0,05% concentration of a test chemical, the cytotoxicity is quantitatively measured as the relative viability of SIRC cells using the MTT assay (4). Decreased cell viability is used to predict potential adverse effects leading to ocular damage.

12. It has been reported that 80% of a solution dropped into the eye of a rabbit is excreted through the conjunctival sac within three to four minutes, while greater than 80% of a solution dropped into the human eye is excreted within one to two minutes (10). The STE test method attempts to approximate these exposure times and makes use of cytotoxicity as an endpoint to assess the extent of damage to SIRC cells following a five-minute exposure to the test chemical.

DEMONSTRATION OF PROFICIENCY

13. Prior to routine use of the STE test method described in this test guideline, laboratories should demonstrate technical proficiency by correctly classifying the eleven substances recommended in Table 1. These substances were selected to represent the full range of responses for serious eye damage or eye irritation based on results of in vivo rabbit eye tests (TG 405) and the UN GHS classification system (1). Other selection criteria included that the substances should be commercially available, that high-quality in vivo reference data should be available, and that high quality in vitro data from the STE test method should be available (3). In situations where a listed substance is unavailable or where justifiable, another substance for which adequate in vivo and in vitro reference data are available could be used provided that the same criteria as described here are used.

Table 1: List of Proficiency Substances

Substance	CASRN	Chemical class[1]	Physical state	*In Vivo UN GHS Cat.*[2]	Solvent in STE test	STE UN GHS Cat.
Benzalkonium chloride (10%, aqueous)	8001-54-5	Onium compound	Liquid	Category 1	Saline	Category 1
Triton X-100 (100%)	9002-93-1	Ether	Liquid	Category 1	Saline	Category 1
Acid Red 92	18472-87-2	Heterocyclic compound; Bromine compound; Chlorine compound	Solid	Category 1	Saline	Category 1
Sodium hydroxide	1310-73-2	Alkali; Inorganic chemical	Solid	Category 1[3]	Saline	Category 1
Butyrolactone	96-48-0	Lactone; Heterocyclic compound	Liquid	Category 2A	Saline	No stand-alone prediction can be made
1-Octanol	111-87-5	Alcohol	Liquid	Category 2A/B[4]	Mineral Oil	No stand-alone prediction can be made
Cyclopentanol	96-41-3	Alcohol; Hydrocarbon, cyclic	Liquid	Category 2A/B[5]	Saline	No stand-alone

						prediction can be made
2-Ethoxyethyl acetate	111-15-9	Alcohol; Ether	Liquid	No Category	Saline	No Category
Dodecane	112-40-3	Hydrocarbon, acyclic	Liquid	No Category	Mineral Oil	No Category
Methyl isobutyl ketone	108-10-1	Ketone	Liquid	No Category	Mineral Oil	No Category
Glycerol	56-81-5	Alcohol	Liquid	No Category	Saline	No Category

Abbreviations: CAS RN = Chemical Abstracts Service Registry Number

[1] Chemical classes were assigned using information obtained from previous NICEATM publications and if not available, using the National Library of Medicine's Medical Subject Headings (MeSH®) (via ChemIDplus® [National Library of Medicine], available at http://chem.sis.nlm.nih.gov/chemidplus/) and structure determinations made by NICEATM.

[2] Based on results from the in vivo rabbit eye test (OECD TG 405) and using the UN GHS (1).

[3] Classification as Cat.1 is based on skin corrosive potential of 100% sodium hydroxide (listed as a proficiency chemical with skin corrosive potential in OECD TG 435) and the criterion for UN GHS category 1 (1).

[4] Classification as 2A or 2B depends on the interpretation of the UN GHS criterion for distinguishing between these two categories, i.e., 2 out of 6 vs 4 out of 6 animals with effects at day 7 necessary to generate a Category 2A classification. The in vivo dataset included 2 studies with 3 animals each. In one study two out of three animals showed effects at day 7 warranting a Cat. 2A classification (11), whereas in the second study all endpoints in all three animals recovered to a score of zero by day 7 warranting a Cat. 2B classification (12).

[5] Classification as 2A or 2B depends on the interpretation of the UN GHS criterion for distinguishing between these two categories, i.e., 1 out of 3 vs 2 out of 3 animals with effects at day 7 necessary to generate a Category 2A classification. The in vivo study included 3 animals. All endpoints apart from corneal opacity and conjunctivae redness in one animal recovered to a score of zero by day 7 or earlier. The one animal that did not fully recover by day 7 had a corneal opacity score of 1 and a conjunctivae redness of 1 (at day 7) that fully recovered at day 14 (11).

PROCEDURE

Preparation of the Cellular Monolayer

14. The rabbit cornea cell line, SIRC should be used for performing the STE test method. It is recommended that SIRC cells are obtained from a well-qualified cell bank, such as American Type Culture Collection CCL60.

15. SIRC cells are cultured at 37°C under 5% CO_2 and humidified atmosphere in a culture flask containing a culture medium comprising Eagle's minimum essential medium (MEM) supplemented with 10% fetal bovine serum (FBS), 2 mM L-glutamine, 50–100 units/mL penicillin and 50–100 µg/mL streptomycin. Cells that have become confluent in the culture flask should be separated using trypsin-ethylenediaminetetraacetic acid solution, with or without the use of a cell scraper. Cells are propagated (e.g. 2 to 3 passages) in a culture flask before being employed for routine testing, and should undergo no more than 25 passages from thawing.

16. Cells ready to be used for the STE test are then prepared at the appropriate density and seeded into 96-well plates. The recommended cell seeding density is 6.0×10^3 cells per well when cells are used four days after seeding, or 3.0×10^3 cells per well when cells are used five days after seeding, at a culture volume of 200 µL. Cells used for the STE test that are seeded in a culture medium at the appropriate density will reach a confluence of more than 80% at the time of testing, i.e., four or five days after seeding.

Application of the Test Chemicals and Control Substances

17. The first choice of solvent for dissolving or suspending test chemicals is physiological saline. If the test chemical demonstrates low solubility or cannot be dissolved or suspended uniformly for at least five minutes in saline, 5% DMSO (CAS#67-68-5) in saline is used as a second choice solvent. For test chemicals that cannot be dissolved or suspended uniformly for at least five minutes in either saline or 5% DMSO in saline, mineral oil (CAS#8042-47-5) is used as a third choice solvent. For highly volatile test chemicals (i.e. vapor pressure over 6 kPa) mineral oil is used as a solvent, provided the test chemical dissolves or forms a stable suspension for at least five minutes in mineral oil.

18. Test chemicals are dissolved or suspended uniformly in the selected solvent at 5% (w/w) concentration and further diluted by serial 10-fold dilution to 0.5% and 0.05% concentration. Each test chemical is to be tested at both 5% and 0.05% concentrations. Cells cultured in the 96-well plate are exposed to 200 µL/well of either a 5% or a 0.05% concentration of the test chemical solution (or suspension), for five minutes at room temperature. Test chemicals (mono-constituent substances or multi-constituent substances or mixtures) are considered as neat substances and diluted or suspended according to the method, regardless of their purity.

19. The culture medium described in paragraph 15 is used as a medium control in each plate of each repetition. Furthermore, cells are to be exposed also to solvent control samples in each plate of each repetition. The solvents listed in paragraph 17 have been confirmed to have no adverse effects on the viability of SIRC cells.

20. In the STE test method, 0.01% Sodium lauryl sulfate (SLS) in saline is to be used as a positive control in each plate of each repetition. In order to calculate cell viability of the positive control, each plate of each repetition has to also include a saline solvent control.

21. A blank is necessary to determine compensation for optical density and should be performed on wells containing culture medium or phosphate buffered saline, but no calcium and magnesium (PBS-) or cells.

22. Each sample (test chemical at 5% and 0.05%, medium control, solvent control, and positive control) should be tested in triplicate in each repetition by exposing the cells to 200 µL of the appropriate test or control chemical for five minutes at room temperature.

23. Benchmark substances are useful for evaluating the ocular irritancy potential of unknown chemicals of a specific chemical or product class, or for evaluating the relative irritancy potential of an ocular irritant within a specific range of irritant responses.

Cell Viability Measurement

24. After exposure, cells are washed twice with 200 µL of PBS and 200 µL of MTT solution (0.5 mg MTT/mL of culture medium) is added. After a two-hour reaction time in an incubator (37˚C, 5% CO_2), the MTT solution is decanted, MTT formazan is extracted by adding 200 µL of 0.04 N hydrochloric acid-isopropanol for 60 minutes in the dark at room temperature, and the absorbance of the MTT formazan solution is measured at 570 nm with a plate reader. Interference of test chemicals with the MTT assay (by colorants or direct MTT reducers) only occurs if significant amount of test chemical is retained in the test system following rinsing after exposure. While this is often the case for the 3D reconstructed human cornea or Reconstructed human epidermis, it is less likely to occur in the 2D cell cultures used for the STE test method. However, because residual material from colorants or direct MTT reducers could interfere with the measurement of optical density, STE users should evaluate such results with caution. If the test results in a Category 1 prediction, then no further actions to address potential interference are needed. Where possible, data should be generated to determine whether such interference is occurring (e.g., conducting an experiment to compare MTT assay OD measurements from test article-treated wells containing SIRC cells in comparison to test article-treated wells containing no cells). If MTT interference is expected to affect the results, alternative cytotoxicity assays (e.g. neutral red) can be used as long as it can be shown to provide similar results as MTT assay, e.g. by testing the proficiency substances in Table 1, and if historical data are available to derive comparable run acceptance criteria (see paragraph 29).

Interpretation of Results and Prediction Model

25. The optical density (OD) values obtained for each test chemical are then used to calculate cell viability relative to the solvent control, which is set at 100%. The relative cell viability is expressed as a percentage and obtained by dividing the OD of test chemical by the OD of the solvent control after subtracting the OD of blank from both values.

$$\text{Cell viability (\%)} = \frac{(OD_{570} \text{ of test chemical}) - (OD_{570} \text{ of blank})}{(OD_{570} \text{ of solvent control}) - (OD_{570} \text{ of blank})} \times 100$$

Similarly, the relative cell viability of each solvent control is expressed as a percentage and obtained by dividing the OD of each solvent control by the OD of the medium control after subtracting the OD of blank from both values.

26. Three independent repetitions, each containing three replicate wells (i.e., n=9), should be performed. The arithmetic mean of the three wells for each test chemical and solvent control in each independent repetition is used to calculate the arithmetic mean of relative cell viability. The final arithmetic mean of the cell viability is calculated from the three independent repetitions.

27. The cell viability cut-off values for identifying test chemicals inducing serious eye damage (UN GHS Category 1) and test chemicals not requiring classification for eye irritation or serious eye damage (UN GHS No Category) are given hereafter.

Table 2: Prediction model of the STE test method

Cell viability		UN GHS Classification	Applicability
At 5%	At 0.05%		
> 70%	> 70%	No Category	Substances and mixtures[1], with the exception of solid chemicals (substances and mixtures) other than surfactants and mixtures composed only of surfactants
≤ 70%	> 70%	No stand-alone prediction can be made[2]	Not applicable
≤ 70%	≤ 70%	Category 1	Substances and mixtures [3]

[1] Mixtures containing test chemicals with vapour pressure higher than 6kPa and that do not either dissolve or form a stable suspension in mineral oil are currently not within the applicability domain of the test method and can generate under-predictions.

[2] No stand-alone prediction can be made from this result in isolation. The result of the STE test should be considered in the context of an IATA (14) for classification purposes.

[3] Based on results obtained mainly with mono-constituent substances, although a limited amount of data also exist on the testing of mixtures. The test method is nevertheless technically applicable to the testing of multi-constituent substances and mixtures. Before use of this Test Guideline on a mixture for generating data for an intended regulatory purpose, it should be considered whether, and if so why, it may provide adequate results for that purpose. Such considerations are not needed, when there is a regulatory requirement for testing of the mixture.

Acceptance Criteria

28. Test results are judged to be acceptable when the following criteria are all satisfied:

a) Optical density of the medium control (exposed to culture medium) should be 0.3 or higher after subtraction of blank optical density.

b) Viability of the solvent control should be 80% or higher relative to the medium control. If multiple solvent controls are used in each repetition, each of those controls should show cell viability greater than 80% to qualify the test chemicals tested with those solvents.

c) The cell viability obtained with the positive control (0.01% SLS) should be within two standard deviations of the historical mean. The upper and lower acceptance boundaries for the positive control should be frequently updated i.e., every three months, or each time an acceptable test is conducted in laboratories where tests are conducted infrequently (i.e., less than once a month). Where a laboratory does not complete a sufficient number of experiments to establish a statistically robust positive control distribution, it is acceptable that the upper and lower acceptance boundaries established by the method developer are used, i.e.,

between 21.1% and 62.3% according to its laboratory historical data, while an internal distribution is built during the first routine tests.

If any of the above criteria a), b) or c) are not met, an additional repetition should be performed.

d) Standard deviation of the final cell viability derived from three independent repetitions should be less than 15% for both 5% and 0.05% concentrations of the test chemical. If the standard deviation is greater than or equal to 15%, the results should not be used and three more repetitions should be performed.

DATA AND REPORTING

Data

29. Data for each individual well (e.g., cell viability values) of each repetition as well as overall mean, standard deviation (SD), and classification are to be reported.

Test Report

30. The test report should include the following information:

Test Chemical and Control Substances

- Mono-constituent substance: chemical identification, such as IUPAC or CAS name(s), CAS registry number(s), SMILES or InChI code, structural formula, and/or other identifiers;
- Multi-constituent substance, UVCB and mixture: Characterization as far as possible by e.g., chemical identity (see above), purity, quantitative occurrence and relevant physicochemical properties (see above) of the constituents, to the extent available;
- Physical state, volatility, pH, LogP, molecular weight, chemical class, and additional relevant physicochemical properties relevant to the conduct of the study, to the extent available;
- Purity, chemical identity of impurities as appropriate and practically feasible, etc;
- Treatment prior to testing, if applicable (e.g., warming, grinding);
- Storage conditions and stability to the extent available;
- Solubility for at least five minutes in a selected solvent (e.g. dissolution or stable suspension).

Test Method Conditions and Procedures

- Name and address of the sponsor, test facility and study director;
- Description of the test method used;
- Cell line used, its source, passage number and confluence of cells used for testing;
- Details of test procedure used;
- Number of repetitions and replicates used;

- Test chemical concentrations used (if different than the ones recommended);
- Justification for choice of solvent for each test chemical;
- Duration of exposure to the test chemical (if different than the one recommended);
- Description of any modifications of the test procedure;
- Description of evaluation and decision criteria used;
- Reference to historical positive control mean and Standard Deviation (SD):
- Demonstration of proficiency of the laboratory in performing the test method (e.g. by testing of proficiency substances) or demonstration of reproducible performance of the test method over time.

Results

- For each test chemical and control substance, and each tested concentration, tabulation should be given for the individual OD values per replicate well, the arithmetic mean OD values for each independent repetition, the % cell viability for each independent repetition, and the final arithmetic mean % cell viability and SD over the three repetitions;
- Results for the medium, solvent and positive control demonstrating suitable study acceptance criteria;
- Description of other effects observed, including retainment of significant amounts of coloured and/or direct MTT reducer test chemical following rinsing after exposure;
- The overall derived classification with reference to the prediction model/decision criteria used.

Discussion of the Results

Conclusions

LITERATURE

1) United Nations UN (2013). Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Fifth revised edition. New York & Geneva: United Nations Publications. ISBN: 978-92-1-117006-1. Available at: [http://www.unece.org/trans/danger/publi/ghs/ghs_rev05/05files_e.html].

2) Scott L, et al. (2010). A proposed Eye Irritation Testing Strategy to Reduce and Replace in vivo Studies Using Bottom-Up and Top-Down Approaches. Toxicol. In Vitro 24, 1-9.

3) Mosmann T. (1983). Rapid Colorimetric Assay for Cellular Growth and Survival: Application to 7 Proliferation and Cytotoxicity Assays. J. Immunol. Methods 65, 55-63.

4) Takahashi Y, et al. (2008). Development of the Short Time Exposure (STE) Test: an In Vitro Eye Irritation Test Using SIRC Cells. Toxicol. In Vitro 22,760-770.

5) Sakaguchi H, et al. (2011). Validation Study of the Short Time Exposure (STE) Test to Assess the Eye Irritation Potential of Chemicals. Toxicol. In Vitro 25,796-809.

6) Kojima H, et al. (2013). Second-Phase Validation of Short Time Exposure Tests for Assessment of Eye Irritation Potency of Chemicals. Toxicol. In Vitro 27, 1855-1869.

7) ICCVAM (2013). Short Time Exposure (STE) Test Method Summary Review Document, NIH. Available at: [http://www.ntp.niehs.nih.gov/iccvam/docs/ocutox_docs/STE-SRD-NICEATM-508.pdf].

8) OECD (2012). Test Guideline for Testing of Chemicals (No.405): Acute Eye Irritation/Corrosion.Organisation for Economic Cooperation and Development, Paris.

9) Saito K, et al. (2015). Predictive Performance of the Short Time Exposure Test for Identifying Eye Irritation Potential of Chemical Mixtures. Toxicol. In Vitro 29, 617–620

10) Mikkelson TJ, Chrai SS and Robinson JR. (1973). Altered Bioavailability of Drugs in the Eye Due to Drug-Protein Interaction. J. Pharm. Sci.1648-1653.

11) ECETOC. (1998). Eye Irritation Reference Chemicals Data Bank. Technical Report (No. 48. (2)), Brussels, Belgium.

12) Gautheron P, et al. (1992). Bovine Corneal Opacity and Permeability Test: an In Vitro Assay of Ocular Irritancy. Fundam Appl Toxicol. 18, 442–449.

13) OECD (2005), Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. Environmental, Health and Safety Publications,Series on Testing and Assessment (No. 34.). Organisation for Economic Cooperation and Development, Paris.

14) OECD (2018). Guidance Document on an Integrated Approaches on Testing and Assessment for Serious Eye Damage and Eye irritation. Series on Testing and Assessment No.263. ENV Publications, Organisation for Economic Cooperation and Development, Paris.

15) Abo T, et al. (2018). Expansion of the applicability domain for highly volatile substances on the Short Time Exposure test method and the predictive performance in assessing eye irritation potential. J. Toxicol. Sci., 43, 407-422

ANNEX I

DEFINITIONS

Accuracy: The closeness of agreement between test method results and accepted reference values. It is a measure of test method performance and one aspect of relevance. The term is often used interchangeably with concordance to mean the proportion of correct outcomes of a test method (13).

Benchmark substance: A substance used as a standard for comparison to a test chemical. A benchmark substance should have the following properties; (i) a consistent and reliable source(s); (ii) structural and functional similarity to the class of substances being tested; (iii) known physical/chemical characteristics; (iv) supporting data on known effects, and (v) known potency in the range of the desired response.

Bottom-Up Approach: A step-wise approach used for a test chemical suspected of not requiring classification for eye irritation or serious eye damage, which starts with the determination of chemicals not requiring classification (negative outcome) from other chemicals (positive outcome)

Chemical: means a substance or mixture.

Eye irritation: Production of change in the eye following the application of a test chemical to the anterior surface of the eye, which are fully reversible within 21 days of application. Interchangeable with "reversible effects on the eye" and with UN GHS Category 2 (1)

False negative rate: The proportion of all positive chemicals falsely identified by a test method as negative. It is one indicator of test method performance.

False positive rate: The proportion of all negative chemicals that are falsely identified by a test method as positive. It is one indicator of test method performance.

Hazard: Inherent property of an agent or situation having the potential to cause adverse effects when an organism, system or (sub) population is exposed to that agent.

Medium control: An untreated replicate containing all components of a test system. This sample is processed with test chemical-treated samples and other control samples to determine whether the solvent interacts with the test system.

Mixture: A mixture or a solution composed of two or more substances in which they do not react (1).

Mono-constituent substance: A substance, defined by its quantitative composition, in which one main constituent is present to at least 80% (w/w).

MTT: 3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide; Thiazolyl blue tetrazolium bromide.

Multi-constituent substance: A substance, defined by its quantitative composition, in which more than one main constituent is present in a concentration ≥ 10% (w/w) and < 80% (w/w). A multi-constituent substance is the result of a manufacturing process. The difference between mixture and multi-constituent substance is that a mixture is obtained by blending of two or more substances without chemical reaction. A multi-constituent substance is the result of a chemical reaction.

OD: Optical Density.

Positive control: A replicate containing all components of a test system and treated with a substance known to induce a positive response. To ensure that variability in the positive control response across time can be assessed, the magnitude of the positive response should not be excessive.

Relevance: Description of relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method (10).

Reliability: Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility and intra-laboratory repeatability (13).

Sensitivity: The proportion of all positive/active chemicals that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results, and is an important consideration in assessing the relevance of a test method (10).

Serious eye damage: Production of tissue damage in the eye, or serious physical decay of vision, following application of a test chemical to the anterior surface of the eye, which is not fully reversible within 21 days of application. Interchangeable with "irreversible effects on the eye" and with UN GHS Category 1 (1).

Solvent/vehicle control: An untreated sample containing all components of a test system, including the solvent or vehicle that is processed with the test chemical-treated and other control samples to establish the baseline response for the samples treated with the test chemical dissolved in the same solvent or vehicle. When tested with a concurrent medium control, this sample also demonstrates whether the solvent or vehicle interacts with the test system.

Specificity: The proportion of all negative/inactive chemicals that are correctly classified by the test. It is a measure of accuracy for a test method that produces categorical results and is an important consideration in assessing the relevance of a test method (13).

Substance: Chemical elements and their compounds in the natural state or obtained by any production process, inducing any additive necessary to preserve the stability of the product and any impurities deriving from the process used, but excluding any solvent which may be separated without affecting the stability of the substance or changing it composition (1).

Surfactant: Also called surface-active agent, this is a chemical such as a detergent, that can reduce the surface tension of a liquid and thus allow it to foam or penetrate solids; it is also known as a wetting agent.

Test chemical: The term "test chemical" is used to refer to what is being tested.

Tiered testing strategy: A stepwise testing strategy where all existing information on a test chemical is reviewed, in a specified order, using a weight of evidence process at each tier to determine if sufficient information is available for a hazard classification decision, prior to progression to the next tier. If the irritancy potential of a test chemical can be assigned based on the existing information, no additional testing is required. If the irritancy potential of a test chemical cannot be assigned based on the existing information, a step-

wise sequential animal testing procedure is performed until an unequivocal classification can be made.

Top-Down Approach: step-wise approach used for a test chemical suspected of causing serious eye damage, which starts with the determination of chemicals inducing serious eye damage (positive outcome) from other chemicals (negative outcome).

United Nations Globally Harmonized System of Classification and Labelling of Chemicals (UN GHS): A system proposing the classification of chemicals (substances and mixtures) according to standardized types and levels of physical, health and environmental hazards, and addressing corresponding communication elements, such as pictograms, signal words, hazard statements, precautionary statements and safety data sheets, so that to convey information on their adverse effects with a view to protect people (including employers, workers, transporters, consumers and emergency responders) and the environment (1).

UN GHS Category 1: See "Serious eye damage".

UN GHS Category 2: See "Eye irritation".

UN GHS No Category: Chemicals that are not classified as UN GHS Category 1 or 2 (2A or 2B).

UVCB: substances of unknown or variable composition, complex reaction products or biological materials.

Section 4
Health effects

Test Guideline No. 442C

In Chemico Skin Sensitisation

Assays addressing the Adverse Outcome Pathway key event on covalent binding to proteins

26 June 2020

OECD Guidelines for the Testing of Chemicals

OECD
BETTER POLICIES FOR BETTER LIVES

OECD/OCDE 442C

Adopted:
18 June 2019
Corrected:
26 June 2020
(Annex 1, MW propyl paraben)

OECD GUIDELINE FOR THE TESTING OF CHEMICALS

Key–Event-Based Test Guideline For In Chemico Skin Sensitisation Assays Addressing The Adverse Outcome Pathway Key Event On Covalent Binding To Proteins

GENERAL INTRODUCTION

Covalent binding to proteins Key Event based Test Guideline.

1. A skin sensitiser refers to a substance that will lead to an allergic response following repeated skin contact as defined by the United Nations Globally Harmonized System of Classification and Labelling of Chemicals (UN GHS) (1). There is general agreement on the key biological events underlying skin sensitisation. The current knowledge of the chemical and biological mechanisms associated with skin sensitisation has been summarised as an Adverse Outcome Pathway (AOP) (2) starting with a molecular initiating event through intermediate events to the adverse effect, namely allergic contact dermatitis. This AOP focuses on chemicals that react with amino-acid residues (i.e. cysteine or lysine) such as organic chemicals. In this instance, the molecular initiating event (i.e. the first key event), is the covalent binding of electrophilic substances to nucleophilic centres in skin proteins. The second key event in this AOP takes place in the keratinocytes and includes inflammatory responses as well as changes in gene expression associated with specific cell signaling pathways such as the antioxidant/electrophile response element (ARE)-dependent pathways. The third key event is the activation of dendritic cells, typically assessed by expression of specific cell surface markers, chemokines and cytokines. The fourth key event is T-cell proliferation.

2. The assessment of skin sensitisation has typically involved the use of laboratory animals. The classical methods that use guinea-pigs, the Guinea Pig Maximisation Test (GPMT) of Magnusson and Kligman and the Buehler Test (OECD TG 406) (11) assess both the induction and elicitation phases of skin sensitisation. The murine tests, such as the LLNA (OECD TG 429) (12) and its three non-radioactive modifications — LLNA:DA (OECD TG 442A) (13), LLNA:BrdU-ELISA, and BrdU-FCM (OECD TG 442B) (14) — all assess the induction response exclusively and have gained acceptance, since they provide an advantage over the guinea pig tests in terms of animal welfare together with an objective measurement of the induction phase of skin sensitisation.

3. Mechanistically-based in chemico and in vitro test methods addressing the first three key events of the skin sensitisation AOP have been adopted for contributing to the evaluation of the skin sensitisation hazard potential of chemicals: the present Test Guideline assesses covalent binding to proteins, addressing the first key event; the OECD

TG 442D assesses keratinocyte activation (15), the second key event and the OECD TG 442E addresses the activation of dendritic cells (16), the third key event of the skin sensitisation AOP. Finally, the fourth key event representing T-cell proliferation is indirectly assessed in the murine Local Lymph Node Assay (LLNA) (12).

Background and principles of the test methods included in the Key Event based Test Guideline

4. This Test Guideline (TG) describes in chemico assays that address mechanisms described under the first key event of the AOP for skin sensitisation, namely covalent binding to proteins (2). The Test Guideline comprises test methods to be used for supporting the discrimination between skin sensitisers and non-sensitisers in accordance with the UN GHS (1). The test methods currently described in this Test Guideline are:

- The Direct Peptide Reactivity Assay (DPRA) (Appendix I), and
- The Amino acid Derivative Reactivity Assay (ADRA) (Appendix II).

5. These two test methods are based on in chemico covalent binding to proteins and are considered to be scientifically valid. The DPRA has been evaluated in a European Union Reference Laboratory for Alternatives to Animal Testing (EURL ECVAM)-lead validation study and subsequent independent peer review by the EURL ECVAM Scientific Advisory Committee (ESAC) (3) (4) (5). The ADRA underwent a validation study coordinated by the Japanese Center for the Validation of Alternative Methods (JaCVAM) (6) (7) (8) (9) followed by an independent peer-review (10).

6. The test methods included in this Test Guideline might differ with regard to the procedures used to generate the data but can each be used to address countries' requirements for test results on protein reactivity, while benefiting from the Mutual Acceptance of Data.

7. The correlation of protein reactivity with skin sensitisation potential is well established (17) (18) (19). Nevertheless, since protein reactivity represents only one key event of the skin sensitisation AOP (2) (20), information generated with test methods developed to address this specific key event may not be sufficient as stand-alone methods to conclude on the presence or absence of skin sensitisation potential of chemicals. Therefore data generated with the test methods described in this Test Guideline are proposed to support the discrimination between skin sensitisers (i.e. UN GHS Category 1) and non-sensitisers when used within Integrated Approaches to Testing and Assessment (IATA), together with other relevant complementary information from in vitro assays addressing other key events of the skin sensitisation AOP as well as non-testing methods, including in silico modeling and read-across from chemical analogues (20). Examples on the use of data generated with these methods within Defined Approaches (DAs) i.e. approaches standardised both in relation to the set of information sources used and in the procedure applied to derive predictions—have been published (20) and can be employed as useful elements within IATA.

8. The test methods described in this Test Guideline do not allow either sub-categorisation of skin sensitisers into subcategories 1A and 1B (21), as defined by UN GHS (1) for authorities implementing these two optional subcategories, or potency prediction for safety assessment decisions. However, depending on the regulatory framework, positive

results generated with these methods may be used on their own to classify a chemical into UN GHS Category 1.

9. Definitions are provided in the Annex. Performance Standards for the assessment of proposed similar or modified in vitro skin sensitisation DPRA and ADRA test methods have been developed (22).

Literature for introduction

(1) United Nations (UN) (2017), Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Seventh revised edition, New York and Geneva, United Nations Publications. Available at: [https://www.unece.org/trans/danger/publi/ghs/ghs_rev07/07files_e0.html]

(2) OECD (2012), Series on Testing and Assessment No. 168. The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. Part 1: Scientific Evidence. Organisation for Economic Cooperation and Development, Paris. Available at:http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=ENV/JM/MONO(2012)10/PART1&docLanguage=En

(3) GF Gerberick, Vassallo JD, Bailey RE, Chaney JG, Morrall SW, Lepoittevin JP (2004), Development of a peptide reactivity assay for screening contact allergens. Toxicol Sci. 81, 332-343.

(4) GF Gerberick, Vassallo JD, Foertsch LM, Price BB, Chaney JG, Lepoittevin JP (2007), Quantification of chemical peptide reactivity for screening contact allergens: a classification tree model approach. Toxicol Sci. 97, 417-427. .

(5) EC EURL-ECVAM (2013), Recommendation on the Direct Peptide Reactivity Assay (DPRA) for the skin sensitisation testing Available at: https://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-recommendations/eurl-ecvam-recommendaion-on-the-directpeptide-reactivity-assay-dpra.

(6) M Fujita, Yamamoto Y, Tahara H, Kasahara T, Jimbo Y, Hioki T (2014), Development of a prediction method for skin sensitisation using novel cysteine and lysinederivatives. J PharmacolToxicol Methods. 70, 94-105.

(7) Y Yamamoto, Tahara H, Usami R, Kasahara T, Jimbo Y, Hioki T, Fujita M.(2015) A novel in chemico method to detect skin sensitisers in highly diluted reactionconditions. J Appl Toxicol. 35, 1348-1360.

(8) M Fujita, Yamamoto Y, Watanabe S, Sugawara T, Wakabayashi K, Tahara K, Horie N, Fujimoto K, Kusakari K, Kurokawa Y, Kawakami T, Kojima K, Kojima H, Ono A, Katsuoka Y, Tanabe H, Yokoyama H and Kasahara T (2019), Cause of and Countermeasures for Oxidation of the Cysteine-Derived Reagent Used in the Amino acid Derivative Reactivity Assay, J. Appl. Toxicology, Feb;39(2):191-208 (doi: 10.1002/jat.3707).

(9) OECD (2019), Draft validation report: Amino acid Derivative Reactivity Assay (ADRA) – JaCVAM Validation Study Report. Series on testing and Assessment n° 304. Organisation for Economic Cooperation and Development, Paris.

(10) OECD (2019), Amino acid Derivative Reactivity Assay (ADRA) – Report of the Peer Review Panel. Series on testing and Assessment n° 305. Organisation for Economic Cooperation and Development, Paris.

(11) OECD (1992), OECD Guidelines for the Testing of Chemicals No. 406. Skin Sensitisation. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

(12) OECD (2010), OECD Guidelines for Chemical Testing No. 429. Skin sensitisation: Local Lymph Node assay. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

(13) OECD (2010), OECD Guidelines for Chemical Testing No. 442A.Skin sensitisation: Local Lymph Node assay: DA. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

(14) OECD (2018), OECD Guidelines for Chemical Testing No. 442B. Skin sensitisation: Local Lymph Node assay: BrdU-ELISA or –FCM. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

(15) OECD (2018), OECD Key Event based test Guideline 442D: In vitro Skin Sensitisation Assays Addressing AOP Key Event on Keratinocyte Activation. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

(16) OECD (2018), OECD Key event based test Guideline 442E: In Vitro Skin Sensitisation Assays Addressing the Key Event on Activation of Dendritic Cells on the Adverse Outcome Pathway for Skin Sensitisation. Organisation for Economic Cooperation and Development, Paris. Available at: [http://www.oecd.org/env/testguidelines].

(17) Landsteiner and Jacobs (1936), Studies on the sensitisation of animals with simple chemical compounds. Journal of Experimental Medicine 64:625-639.

(18) Dupuis and Benezra (1982), Allergic contact dermatitis to simple chemicals: a molecular approach. New York & Basel: Marcel Dekker Inc.

(19) JP Lepoittevin, Basketter DA, Goossens A, Karlberg AT (1998), Allergic contact dermatitis: the molecular basis, Springer, Berlin (doi: 10.1007/978-3-642-80331-4).

(20) OECD (2016), Series on Testing & Assessment No. 256: Guidance Document On The Reporting Of Defined Approaches And Individual Information Sources To Be Used Within Integrated Approaches To Testing And Assessment (IATA) For Skin Sensitisation, Annex 1 and Annex 2. ENV/JM/HA(2016)29. Organisation for Economic Cooperation and Development, Paris. Available at: [https://community.oecd.org/community/iatass].

(21) B Wareing, Urbisch D, Kolle SN, Honarvar N, Sauer UG, Mehling A, Landsiedel R(2017) Prediction of skin sensitization potency sub-categories using peptide reactivity data, Toxicol In Vitro Dec;45(Pt 1):134-145 (doi: 10.1016/j.tiv.2017.08.015).

(22) OECD (2019), Performance Standards for the assessment of proposed similar or modified in vitro skin sensitisation DPRA and ADRA test methods, Series on Testing & Assessment No. 303, Organisation for Economic Cooperation and Development, Paris.

ANNEX - DEFINITIONS

Accuracy: The closeness of agreement between test method results and accepted reference values. It is a measure of test method performance and one aspect of relevance. The term is often used interchangeably with concordance to mean the proportion of correct outcomes of a test method (1).

(Formula shown below.)

ADRA: Amino acid Derivative Reactivity Assay

AOP (Adverse Outcome Pathway): sequence of events from the chemical structure of a target chemical or group of similar chemicals through the molecular initiating event to an in vivo outcome of interest (2).

Calculation

Calculating depletion of either NAC or NAL

Depletion is calculated as follows:

Percent depletion of either NAC or NAL = {1- (NAC or NAL peak area in replicate injection ÷ mean NAC or NAL peak area in reference control C)} × 100

Calculating predictive capacity

There are several terms that are commonly used along with the description of sensitivity, specificity and accuracy. They are true positive (TP), true negative (TN), false negative (FN), and false positive (FP).

Sensitivity, specificity and accuracy are described in terms of TP, TN, FN, and FP.

Sensitivity: Number of true positives ÷ Number of all positive chemicals, TP ÷ (TP + FN)

Specificity: Number of true negatives ÷ Number of all negative chemicals, TN ÷ (TN + FP)

Accuracy: Number of correct predictions ÷ Number of all predictions, (TN + TP) ÷ (TN+TP+FN+FP)

Calibration curve: The relationship between the experimental response value and the analytical concentration (also called standard curve) of a known substance.

Coefficient of variation: a measure of variability that is calculated for a group of replicate data by dividing the standard deviation by the mean. It can be multiplied by 100 for expression as a percentage.

Defined Approach (DA): a DA consists of a fixed data interpretation procedure (e.g. statistical, mathematical models) applied to data (e.g. in silico predictions, in chemico, in vitro data) generated with a defined set of information sources to derive a prediction.

DPRA: Direct Peptide Reactivity Assay

EDTA: Ethylenediaminetetraacetic acid

EURL ECVAM: the European Union Reference Laboratory for Alternatives to Animal Testing

Hazard: Inherent property of an agent or situation having the potential to cause adverse effects when an organism, system or (sub) population is exposed to that agent.

IATA (Integrated Approach to Testing and Assessment): A structured approach used for hazard identification (potential), hazard characterisation (potency), and/or safety assessment (potential/potency and exposure) of a chemical or group of chemicals, which strategically integrates and weights all relevant data to inform regulatory decision regarding potential hazards, risks, and the need for further targeted and therefore minimal testing.

JaCVAM: Japanese Center for the Validation of Alternative Methods

LLNA: murine Local Lymph Node Assay issued as OECD TG 429 in 2010

Molecular Initiating Event: Chemical-induced perturbation of a biological system at the molecular level identified to be the starting event in the adverse outcome pathway.

Mixture: A solid or liquid comprising two or more substances which do not react chemically. (3)

Mono-constituent substance: A substance, defined by its quantitative composition, in which one main constituent comprises at least 80% (w/w) of the whole.

Multi-constituent substance: A substance, defined by its quantitative composition, in which two or more main constituents are present in concentrations ≥ 10% (w/w) and < 80% (w/w). Multi-constituent substances are the result of a manufacturing process. The difference between a mixture and a multi-constituent substance is that a mixture comprises two or more substances which do not react chemically, whereas a multi-constituent substance comprises two or more substances that do react chemically.

NAC: N-(2-(1-naphthyl)acetyl)-L-cysteine (4) (5) (6)

NAL: α-N-(2-(1-naphthyl)acetyl)-L-lysine (4) (5) (6)

Positive control: A replicate containing all components of a test system and treated with a substance known to induce a positive response. To ensure that variability in the positive control response across time can be assessed, the magnitude of the positive response should not be excessive.

Pre-haptens: chemicals which become sensitisers through abiotic transformation

Pro-haptens: chemicals requiring enzymatic activation to exert skin sensitisation potential

Reference control: An untreated sample containing all components of a test system, including the solvent or vehicle that is processed with the test chemical treated and other control samples to establish the baseline response for the samples treated with the test chemical dissolved in the same solvent or vehicle. When tested with a concurrent negative control, this sample also demonstrates whether the solvent or vehicle interacts with the test system.

Relevance: Description of relationship of the test to the effect of interest and whether it is meaningful and useful for a particular purpose. It is the extent to which the test correctly measures or predicts the biological effect of interest. Relevance incorporates consideration of the accuracy (concordance) of a test method. (1)

Reliability: Measures of the extent that a test method can be performed reproducibly within and between laboratories over time, when performed using the same protocol. It is assessed by calculating intra- and inter-laboratory reproducibility and intra-laboratory repeatability. (1)

Reproducibility: The concordance of results obtained from testing the same substance using the same test protocol (see reliability). (1)

Sensitivity: The proportion of all positive/active chemicals that are correctly classified by the test method. It is a measure of accuracy for a test method that produces categorical results and is an important consideration in assessing the relevance of a test method. (1) (Formula shown below.)

Specificity: The proportion of all negative/inactive chemicals that are correctly classified by the test method. It is a measure of accuracy for a test method that produces categorical results and is an important consideration in assessing the relevance of a test method. (1) (Formula shown below.)

Substance: Chemical elements and their compounds in the natural state or resulting from a manufacturing process, including any additive necessary to preserve the stability of the product and any impurities deriving from the process, but excluding solvents that may be separated without affecting the stability of the substance or changing its composition (3).

System suitability: Determination of instrument performance (e.g. sensitivity) by analysis of a reference standard prior to running the analytical batch (7).

Test chemical: The term test chemical is used to refer to the substance being tested.

TFA: Trifluoroacetic acid

United Nations Globally Harmonized System of Classification and Labelling of Chemicals (UN GHS): A system proposing the classification of chemicals (substances and mixtures) according to standardised types and levels of physical, health and environmental hazards, and addressing corresponding communication elements, such as pictograms, signal words, hazard statements, precautionary statements and safety data sheets, so that to convey information on their adverse effects with a view to protect people (including employers, workers, transporters, consumers and emergency responders) and the environment (3).

UVCB: substances of unknown or variable composition, complex reaction products or biological materials.

Valid test method: A test method considered to have sufficient relevance and reliability for a specific purpose and which is based on scientifically sound principles. A test method is never valid in an absolute sense, but only in relation to a defined purpose (1).

Literature for definitions

(1) OECD (2005), Guidance Document on the Validation and International Acceptance of New or Updated Test Methods for Hazard Assessment. OECD Series on Testing and Assessment, No. 34. Organisation for Economic Cooperation and Development, Paris, France.

(2) OECD (2012), The Adverse Outcome Pathway for Skin Sensitisation Initiated by Covalent Binding to Proteins. Part 1: Scientific Evidence. Series on Testing and Assessment No. 168, OECD, Paris.

(3) United Nations (UN) (2013), Globally Harmonized System of Classification and Labelling of Chemicals (GHS). Fifth revised edition, UN New York and Geneva, 2013. Available at: http://www.unece.org/trans/danger/publi/ghs/ghs_rev05/05files_e.html

(4) M Fujita, Yamamoto Y, Tahara H, Kasahara T, Jimbo Y and Hioki T (2014), Development of a prediction method for skin sensitisation using novel cysteine and lysine derivatives, Journal of pharmacological and toxicological methods, 70:94-105.

(5) Y Yamamoto, Tahara H, Usami R, Kasahara T, Jimbo Y, Hioki T and Fujita M (2015), A novel in chemico method to detect skin sensitisers in highly diluted reaction conditions, Journal of Applied Toxicology, 35(11):1348-60, (doi: 10.1002/jat.3139).

(6) M Fujita, Yamamoto Y, Watanabe S, Sugawara T, Wakabayashi K, Tahara K, Horie N, Fujimoto K, Kusakari K, Kurokawa Y, Kawakami T, Kojima K, Kojima H, Ono A, Katsuoka Y, Tanabe H, Yokoyama H and Kasahara T (2019), Cause of and Countermeasures for Oxidation of the Cysteine-Derived Reagent Used in the Amino acid Derivative Reactivity Assay, J. Appl. Toxicology, Feb;39(2):191-208 (doi: 10.1002/jat.3707).

(7) FDA (Food and Drug Administration) (2001), Guidance for Industry: Bioanalytical Method Validation 22pp. Accessible at:www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidance/ucm070 107.pdf - 138 (23)

APPENDIX I

In Chemico Skin Sensitisation: Direct Peptide Reactivity Assay (DPRA)

INITIAL CONSIDERATIONS, APPLICABILITY AND LIMITATIONS

1. The DPRA is proposed to address the molecular initiating event of the skin sensitisation AOP, namely protein reactivity, by quantifying the reactivity of test chemicals towards model synthetic peptides containing either lysine or cysteine (1). Cysteine and lysine percent peptide depletion values are then used to categorise a substance in one of four classes of reactivity for supporting the discrimination between skin sensitisers and non-sensitisers (2).

2. The DPRA test method proved to be transferable to laboratories experienced in high-performance liquid chromatography (HPLC) analysis. The level of reproducibility in predictions that can be expected from the test method is in the order of 85% within laboratories and 80% between laboratories (3). Results generated in the validation study (4) and published studies (5) overall indicate that the accuracy of the DPRA in discriminating sensitisers (i.e. UN GHS Cat. 1) from non-sensitisers is 80% (N=157) with a sensitivity of 80% (88/109) and specificity of 77% (37/48) when compared to LLNA results. The DPRA is more likely to under predict chemicals showing a low to moderate skin sensitisation potency (i.e. UN GHS subcategory 1B) than chemicals showing a high skin sensitisation potency (i.e. UN GHS subcategory 1A) (4) (5). However, the accuracy values given here for the DPRA as a stand-alone test method are only indicative since the test method should be considered in combination with other sources of information in the context of an IATA or a DA and in accordance with the provisions of paragraphs 7 and 8 in the General introduction. Furthermore when evaluating non-animal methods for skin sensitisation, it should be kept in mind that the LLNA test as well as other animal tests may not fully reflect the situation in the species of interest, i.e. humans. On the basis of the overall data available, the DPRA was shown to be applicable to test chemicals covering a variety of organic functional groups, reaction mechanisms, skin sensitisation potency (as determined in *in vivo* studies) and physico-chemical properties (1) (2) (3) (5). Taken together, this information indicates the usefulness of the DPRA to contribute to the identification of skin sensitisation hazard.

3. The term "test chemical" is used in this Test Guideline to refer to what is being tested[1] and is not related to the applicability of the DPRA to the testing of substances and/or mixtures. This test method is not applicable for the testing of metal compounds since they are known to react with proteins with mechanisms other than covalent binding. A test chemical should be soluble in an appropriate solvent at a final concentration of 100 mM (see paragraph 10). However, test chemicals that are not soluble at this concentration may still be tested at lower soluble concentrations. In such a case, a positive result could still be used to support the identification of the test chemical as a skin sensitiser but no firm

[1] In June 2013, the Joint Meeting agreed that where possible, a more consistent use of the term “test chemical” describing what is being tested should now be applied in new and updated Test Guidelines.

conclusion on the lack of reactivity should be drawn from a negative result. Limited information is currently available on the applicability of the DPRA to mixtures of known composition (4) (5). The DPRA is nevertheless considered to be technically applicable to the testing of multi-constituent substances and mixtures of known composition (see paragraph 4 and 10). When considering testing of mixtures, difficult-to-test chemicals (e.g. unstable), or test chemicals not clearly within the applicability domain described in this Appendix of the Test Guideline, upfront consideration should be given to whether the results of such testing will yield results that are meaningful scientifically. The current prediction model cannot be used for complex mixtures of unknown composition or for substances of unknown or variable composition, complex reaction products or biological materials (i.e. UVCB substances) due to the defined molar ratio of test chemical and peptide. For this purpose a new prediction model based on a gravimetric approach will need to be developed. In cases where evidence can be demonstrated on the non-applicability of the test method to other specific categories of chemicals, the test method should not be used for those specific categories of chemicals.

4. The test method described in this Appendix of the Test Guideline is an *in chemico* method that does not encompass a metabolic system. Chemicals that require enzymatic bioactivation to exert their skin sensitisation potential (i.e. pro-haptens) cannot be detected by the test method. Chemicals that become sensitisers after abiotic transformation (i.e. pre-haptens) are reported to be in most cases correctly detected by the test method (4) (9) (10). In the light of the above, negative results obtained with the test method should be interpreted in the context of the stated limitations and in the connection with other information sources within the framework of an IATA or a DA. Test chemicals that do not covalently bind to the peptide but promote its oxidation (i.e. cysteine dimerisation) could lead to a potential over estimation of peptide depletion, resulting in possible false positive predictions and/or assignment to a higher reactivity class (see paragraphs 21 and 22).

5. As described, the DPRA assay supports the discrimination between skin sensitisers and non-sensitisers. However, it may also potentially contribute to the assessment of sensitising potency (6) (11) when used in integrated approaches such as IATA or DA (12). However further work, preferably based on human data, is required to determine how DPRA results may possibly inform potency assessment.

PRINCIPLE OF THE TEST

6. The DPRA is an *in chemico* method which quantifies the remaining concentration of cysteine- or lysine-containing peptide following 24 hours incubation with the test chemical at 22.5-30°C. The synthetic peptides contain phenylalanine to aid in the detection. Relative peptide concentration is measured by high-performance liquid chromatography (HPLC) with gradient elution and UV detection at 220 nm. Cysteine- and lysine peptide percent depletion values are then calculated and used in a prediction model (see paragraph 21) which allows assigning the test chemical to one of four reactivity classes used to support the discrimination between sensitisers and non-sensitisers.

7. Prior to routine use of the method described in this Appenix, laboratories should demonstrate technical proficiency, using the ten proficiency substances listed in Annex 1.

PROCEDURE

8. This test method is based on the DPRA DB-ALM protocol nº 154 (7) which represents the protocol used for the EURL ECVAM-coordinated validation study. It is recommended that this protocol is used when implementing and using the method in the laboratory. The following is a description of the main components and procedures for the DPRA. If an alternative HPLC set-up is used, its equivalence to the validated set-up described in the DB-ALM protocol should be demonstrated (e.g. by testing the proficiency substances in Annex 1).

Preparation of the cysteine or lysine-containing peptides

9. Stock solutions of cysteine (Ac-RFAA**C**AA-COOH) and lysine (Ac-RFAA**K**AA-COOH) containing synthetic peptides of purity higher than 85% and preferably > 90%, should be freshly prepared just before their incubation with the test chemical. The final concentration of the cysteine peptide should be 0.667 mM in pH 7.5 phosphate buffer whereas the final concentration of the lysine peptide should be 0.667 mM in pH 10.2 ammonium acetate buffer. The HPLC run sequence should be set up in order to keep the HPLC analysis time less than 30 hours. For the HPLC set up used in the validation study and described in this test method, up to 26 analysis samples (which include the test chemical, the positive control and the appropriate number of solvent controls based on the number of individual solvents used in the test, each tested in triplicate), can be accommodated in a single HPLC run. All of the replicates analysed in the same run should use the identical cysteine and lysine peptide stock solutions. It is recommended to prove individual peptide batches for proper solubility prior to their use.

Preparation of the test chemical

10. Solubility of the test chemical in an appropriate solvent should be assessed before performing the assay following the solubilisation procedure described in the DPRA DB-ALM protocol (7). An appropriate solvent will dissolve the test chemical completely. Since in the DPRA the test chemical is incubated in large excess with either the cysteine or the lysine peptides, visual inspection of the forming of a clear solution is considered sufficient to ascertain that the test chemical (and all of its components in the case of testing a multi-constituent substance or a mixture) is dissolved. Suitable solvents are, acetonitrile, water, 1:1 mixture water:acetonitrile, isopropanol, acetone or 1:1 mixture acetone:acetonitrile. Other solvents can be used as long as they do not have an impact on the stability of the peptide as monitored with reference controls C (i.e. samples constituted by the peptide alone dissolved in the appropriate solvent; see Annex 2). If the test chemical is not soluble in any of the solvents mentioned above, DMSO can be used as a last resort and in minimal amounts. It is important to note that DMSO may lead to peptide dimerisation and as a result, it may be more difficult to meet the acceptance criteria. If DMSO is chosen, attempts should be made to first solubilise the test chemical in 300 µL of DMSO and dilute the resulting solution with 2700 µL of acetonitrile. If the test chemical is not soluble in this mixture, attempts should be made to solubilise the same amount of test chemicals in 1500 µL of DMSO and dilute the resulting solution with 1500 µL of acetonitrile. The test chemical should be pre-weighed into glass vials and dissolved immediately before testing in an appropriate solvent to prepare a 100 mM solution. For mixtures and multi-constituent substances of known composition, a single purity should be determined by the sum of the proportion of its constituents (excluding water), and a single apparent molecular weight should be determined by considering the individual molecular weights of each component

in the mixture (excluding water) and their individual proportions. The resulting purity and apparent molecular weight should then be used to calculate the weight of test chemical necessary to prepare a 100 mM solution. For polymers for which a predominant molecular weight cannot be determined, the molecular weight of the monomer (or the apparent molecular weight of the various monomers constituting the polymer) may be considered to prepare a 100 mM solution. However, when testing mixtures, multi-constituent substances or polymers of known composition, it should be considered to also test the neat chemical. For liquids, the neat chemical should be tested as such without any prior dilution by incubating it at 1:10 and 1:50 ratio with the cysteine and lysine peptides, respectively. For solids, the test chemical should be dissolved to its maximum soluble concentration in the same solvent used to prepare the apparent 100 mM solution. It should then be tested as such without any further dilution by incubating it at 1:10 and 1:50 ratio with the cysteine and lysine peptides, respectively. Concordant results (reactive or non-reactive) between the apparent 100 mM solution and the neat chemical should allow for a firm conclusion on the result.

Preparation of the positive control, reference controls and coelution controls

11. Cinnamic aldehyde (CAS 104-55-2; ≥95% food-grade purity) should be used as positive control (PC) at a concentration of 100 mM in acetonitrile. Other suitable positive controls providing mid-range depletion values may be used if historical data are available to derive comparable run acceptance criteria. In addition reference controls (i.e. samples containing only the peptide dissolved in the appropriate solvent) should also be included in the HPLC run sequence and these are used to verify the HPLC system suitability prior to the analysis (reference controls A), the stability of the reference controls over time (reference control B) and to verify that the solvent used to dissolve the test chemical does not impact the percent peptide depletion (reference control C) (see Annex 2). The appropriate reference control for each substance is used to calculate the percent peptide depletion for that substance (see paragraph 18). In addition, a co-elution control constituted by the test chemical alone for each of the test chemicals analysed should be included in the run sequence to detect possible co-elution of the test chemical with either the lysine or the cysteine peptide.

Incubation of the test chemical with the cysteine and lysine peptide solutions

12. Cysteine and lysine peptide solutions should be incubated in glass autosampler vials with the test chemical at 1:10 and 1:50 ratio respectively. If a precipitate is observed immediately upon addition of the test chemical solution to the peptide solution, due to low aqueous solubility of the test chemical, one cannot be sure how much test chemical remained in the solution to react with the peptide. Therefore, in such a case, a positive result could still be used, but a negative result is uncertain and should be interpreted with due care (see also provisions in paragraph 10 for the testing of chemicals not soluble up to a concentration of 100 mM). The reaction solution should be left in the dark at 22.5-30°C for 24±2 hours before running the HPLC analysis. Each test chemical should be analysed in triplicate for both peptides. Samples have to be visually inspected prior to HPLC analysis. If a precipitate or phase separation is observed, samples may be centrifuged at low speed (100-400xg) to force precipitate to the bottom of the vial as a precaution since large amounts of precipitate may clog the HPLC tubing or columns. If a precipitation or phase separation is observed after the incubation period, peptide depletion may be underestimated

and a conclusion on the lack of reactivity cannot be drawn with sufficient confidence in case of a negative result.

Preparation of the HPLC standard calibration curve

13. A standard calibration curve should be generated for both the cysteine and the lysine peptides. Peptide standards should be prepared in a solution of 20% or 25% acetonitrile:buffer using phosphate buffer (pH 7.5) for the cysteine peptide and ammonium acetate buffer (pH 10.2) for the lysine peptide. Using serial dilution standards of the peptide stock solution (0.667 mM), 6 calibration solutions should be prepared to cover the range from 0.534 to 0.0167 mM. A blank of the dilution buffer should also be included in the standard calibration curve. Suitable calibration curves should have an $r^2>0.99$.

HPLC preparation and analysis

14. The suitability of the HPLC system should be verified before conducting the analysis. Peptide depletion is monitored by HPLC coupled with an UV detector (photodiode array detector or fixed wavelength absorbance detector with 220 nm signal). The appropriate column is installed in the HPLC system. The HPLC set-up described in the validated protocol uses a Zorbax SB-C-18 2.1 mm x 100 mm x 3.5 micron as preferred column. With this reversed-phase HPLC column, the entire system should be equilibrated at 30°C with 50% phase A (0.1% (v/v) trifluoroacetic acid in water) and 50% phase B (0.085% (v/v) trifluoroacetic acid in acetonitrile) for at least 2 hours before running. The HPLC analysis should be performed using a flow rate of 0.35 mL/min and a linear gradient from 10% to 25% acetonitrile over 10 minutes, followed by a rapid increase to 90% acetonitrile to remove other materials. Equal volumes of each standard, sample and control should be injected. The column should be re-equilibrated under initial conditions for 7 minutes between injections. If a different reversed-phase HPLC column is used, the set-up parameters described above may need to be adjusted to guarantee an appropriate elution and integration of the cysteine and lysine peptides, including the injection volume, which may vary according to the system used (typically in the range from 3-10 µL). Importantly, if an alternative HPLC set-up is used, its equivalence to the validated set-up described above should be demonstrated (e.g. by testing the proficiency substances in Annex 1). Absorbance is monitored at 220 nm. If a photodiode array detector is used, absorbance at 258 nm should also be recorded. It should be noted that some supplies of acetonitrile could have a negative impact on peptide stability and this has to be assessed when a new batch of acetonitrile is used. The ratio of the 220 peak area and the 258 peak area can be used as an indicator of co-elution. For each sample a ratio in the range of 90%<mean[2] area ratio of control samples<100% would give a good indication that co-elution has not occurred.

15. There may be test chemicals which could promote the oxidation of the cysteine peptide. The peak of the dimerised cysteine peptide may be visually monitored. If dimerisation appears to have occurred, this should be noted as percent peptide depletion may be over-estimated leading to false positive predictions and/or assignment to a higher reactivity class (see paragraphs 21 and 22).

16. The HPLC analysis should be timed to assure that the injection of the first sample starts 22 to 26 hours after the test chemical was mixed with the peptide solution. The HPLC run sequence should be set up in order to keep the HPLC analysis time less than 30 hours. For the HPLC set up used in the validation study and described in this test method, up to

[2] For mean it is meant arithmetic mean throughout the document.

26 analysis samples can be accommodated in a single HPLC run (see also paragraph 9). An example of HPLC analysis sequence is provided in Annex 2.

DATA AND REPORTING

Data evaluation

17. The concentration of cysteine or lysine peptide is photometrically determined at 220 nm in each sample by measuring the peak area (area under the curve, AUC) of the appropriate peaks and by calculating the concentration of peptide using the linear calibration curve derived from the standards.

18. The percent peptide depletion is determined in each sample by measuring the peak area and dividing it by the mean peak area of the relevant reference controls C (see Annex 2) according to the formula described below.

$$\text{Percent peptide depletion} = \left[1 - \left(\frac{\text{Peptide peak area in replicate injection}}{\text{Mean peptide peak area in reference controls } C}\right)\right] \times 100$$

Acceptance criteria

19. The following criteria should be met for a run to be considered valid:

a) the standard calibration curve should have an r^2>0.99,

b) the mean percent peptide depletion value of the three replicates for the positive control cinnamic aldehyde should be between 60.8% and 100% for the cysteine peptide and between 40.2% and 69.0% for the lysine peptide (for other positive controls a reference range needs to be established) and the maximum standard deviation (SD) for the positive control replicates should be <14.9% for the percent cysteine depletion and <11.6% for the percent lysine depletion and

c) the mean peptide concentration of reference controls A should be 0.50±0.05 mM and the coefficient of variation (CV) of peptide peak areas for the nine reference controls B and C in acetonitrile should be <15.0%.

If one or more of these criteria is not met the run should be repeated.

20. The following criteria should be met for a test chemical's results to be considered valid:

a) the maximum standard deviation for the test chemical replicates should be <14.9% for the percent cysteine depletion and <11.6% for the percent lysine depletion,

b) the mean peptide concentration of the three reference controls C in the appropriate solvent should be 0.50±0.05 mM.

If these criteria are not met the data should be rejected and the run should be repeated for that specific test chemical.

Prediction model

21. The mean percent cysteine and percent lysine depletion value is calculated for each test chemical. Negative depletion is considered as "0" when calculating the mean. By using the cysteine 1:10/lysine 1:50 prediction model shown in Table 1, the threshold of 6.38% average peptide depletion should be used to support the discrimination between skin sensitisers and non-sensitisers in the framework of an IATA or DA. Application of the prediction model for assigning a test chemical to a reactivity class (i.e. low, moderate and high reactivity) may perhaps prove useful to inform potency assessment within the framework of an IATA or DA.

Table 1: Cysteine 1:10/lysine 1:50 prediction model[1]

Mean of cysteine and lysine % depletion	Reactivity Class	DPRA Prediction[2]
0% ≤ mean % depletion ≤ 6.38%	No or minimal reactivity	Negative
6.38% < mean % depletion ≤ 22.62%	Low reactivity	Positive
22.62% < mean % depletion ≤ 42.47%	Moderate reactivity	
42.47% < mean % depletion ≤ 100%	High reactivity	

[1] The numbers refer to statistically generated threshold values and are not related to the precision of the measurement.
[2] A DPRA prediction should be considered in the framework of an IATA and in accordance with the provisions of paragraphs 2 and 4.

22. There might be cases where the test chemical (the substance or one or several of the components of a multi-constituent substance or a mixture) absorbs significantly at 220 nm and has the same retention time of the peptide (co-elution). Co-elution may be resolved by slightly adjusting the HPLC set-up in order to further separate the elution time of the test chemical and the peptide. If an alternative HPLC set-up is used to try to resolve co-elution, its equivalence to the validated set-up should be demonstrated (e.g. by testing the proficiency substances in Annex 1). When co-elution occurs the peak of the peptide cannot be integrated and the calculation of the percent peptide depletion is not possible. If co-elution of such test chemicals occurs with both the cysteine and the lysine peptides then the analysis should be reported as "inconclusive". In cases where co-elution occurs only with the lysine peptide, then the cysteine 1:10 prediction model reported in Table 2 can be used.

Table 2: Cysteine 1:10 prediction model[1]

Cysteine (Cys) % depletion	Reactivity class	DPRA prediction[2]
0% ≤ Cys % depletion ≤ 13.89%	No or minimal reactivity	Negative
13.89% < Cys % depletion ≤ 23.09%	Low reactivity	Positive
23.09% < Cys % depletion ≤ 98.24%	Moderate reactivity	
98.24% < Cys % depletion ≤ 100%	High reactivity	

[1] The numbers refer to statistically generated threshold values and are not related to the precision of the measurement.
[2] A DPRA prediction should be considered in the framework of an IATA and in accordance with the provisions of paragraphs 2 and 4.

23. There might be other cases where the overlap in retention time between the test chemical and either of the peptides is incomplete. In such cases percent peptide depletion values can be estimated and used in the cysteine 1:10/lysine 1:50 prediction model, however assignment of the test chemical to a reactivity class cannot be made with accuracy.

24. A single HPLC analysis for both the cysteine and the lysine peptide should be sufficient for a test chemical when the result is unequivocal. However, in cases of results close to the threshold used to discriminate between positive and negative results (i.e. borderline results), additional testing may be necessary. If situations where the mean percent depletion falls in the range of 3% to 10% for the cysteine 1:10/lysine 1:50 prediction model or the cysteine percent depletion falls in the range of 9% to 17% for the cysteine 1:10 prediction model, a second run may be considered, as well as a third one in case of discordant results between the first two runs.

Test report

25. The test report should include the following information

Test chemical

- Mono-constituent substance
    - Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers;
    - Physical appearance, water solubility, molecular weight, and additional relevant physicochemical properties, to the extent available;
    - Purity, chemical identity of impurities as appropriate and practically feasible, etc;
    - Treatment prior to testing, if applicable (e.g. warming, grinding);
    - Concentration(s) tested;
    - Storage conditions and stability to the extent available.
- Multi-constituent substance, UVCB and mixture:
    - Characterisation as far as possible by e.g. chemical identity (see above), purity, quantitative occurrence and relevant physicochemical properties (see above) of the constituents, to the extent available;
    - Physical appearance, water solubility and additional relevant physicochemical properties, to the extent available;
    - Molecular weight or apparent molecular weight in case of mixtures/polymers of known compositions or other information relevant for the conduct of the study;
    - Treatment prior to testing, if applicable (e.g. warming, grinding);
    - Concentration(s) tested;
    - Storage conditions and stability to the extent available.

Controls

- Positive control
    - Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers;
    - Physical appearance, water solubility, molecular weight, and additional relevant physicochemical properties, to the extent available;
    - Purity, chemical identity of impurities as appropriate and practically feasible, etc;
    - Treatment prior to testing, if applicable (e.g. warming, grinding);
    - Concentration(s) tested;
    - Storage conditions and stability to the extent available;
    - Reference to historical positive control results demonstrating suitable run acceptance criteria, if applicable.
- Solvent/vehicle
    - Solvent/vehicle used and ratio of its constituents, if applicable;
    - Chemical identification(s), such as IUPAC or CAS name(s), CAS number(s), and/or other identifiers;
    - Purity, chemical identity of impurities as appropriate and practically feasible, etc;
    - Physical appearance, molecular weight, and additional relevant physicochemical properties in the case other solvents/vehicles than those mentioned in the test method are used and to the extent available;
    - Storage conditions and stability to the extent available;
    - Justification for choice of solvent for each test chemical;
    - For acetonitrile, results of test of impact on peptide stability.

Preparation of peptides, positive control and test chemical

- Characterisation of peptide solutions (supplier, lot, exact weight of peptide, volume added for the stock solution);
- Characterisation of positive control solution (exact weight of positive control substance, volume added for the test solution);
- Characterisation of test chemical solutions (exact weight of test chemical, volume added for the test solution).

HPLC instrument setting and analysis

- Type of HPLC instrument, HPLC and guard columns, detector, autosampler;
- Parameters relevant for the HPLC analysis such as column temperature, injection volumes, flow rate and gradient.

System suitability

- Peptide peak area at 220 nm of each standard and reference control A replicate;
- Linear calibration curve graphically represented and the r^2 reported;
- Peptide concentration of each reference control A replicate;
- Mean peptide concentration (mM) of the three reference controls A, SD and CV;
- Peptide concentration of reference controls A and C.

Analysis sequence

- For reference controls:
    - Peptide peak area at 220 nm of each B and C replicate;
    - Mean peptide peak area at 220 nm of the nine reference controls B and C in acetonitrile, SD an CV (for stability of reference controls over analysis time);
    - For each solvent used, the mean peptide peak area at 220 nm of the three appropriate reference controls C (for the calculation of percent peptide depletion);
    - For each solvent used, the peptide concentration (mM) of the three appropriate reference controls C;
    - For each solvent used, the mean peptide concentration (mM) of the three appropriate reference controls C, SD and CV.
- For positive control:
    - Peptide peak area at 220 nm of each replicate;
    - Percent peptide depletion of each replicate;
    - Mean percent peptide depletion of the three replicates, SD and CV.
- For each test chemical:
    - Appearance of precipitate in the reaction mixture at the end of the incubation time, if observed. If precipitate was re-solubilised or centrifuged;
    - Presence of co-elution;
    - Description of any other relevant observations, if applicable;
    - Peptide peak area at 220 nm of each replicate;
    - Percent peptide depletion of each replicate;
    - Mean of percent peptide depletion of the three replicate, SD and CV;
    - Mean of percent cysteine and percent lysine depletion values;
    - Prediction model used and DPRA prediction.

Proficiency testing

- If applicable, the procedure used to demonstrate proficiency of the laboratory in performing the test method (e.g. by testing of proficiency substances) or to demonstrate reproducible performance of the test method over time.

Discussion of the results

- Discussion of the results obtained with the DPRA test method;
- Discussion of the test method results in the context of an IATA if other relevant information is available.

Conclusion

LITERATURE FOR APPENDIX I

(1) Gerberick *et al.* (2004), Development of a peptide reactivity assay for screening contact allergens. Toxicological Sciences 81:332-343.

(2) Gerberick *et al.* (2007), Quantification of chemical peptide reactivity for screening contact allergens: A classification tree model approach. Toxicological Sciences 97:417-427.

(3) EC EURL-ECVAM (2013), Recommendation on the Direct Peptide Reactivity Assay (DPRA) for skin sensitisation testing. Available at: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-recommendations/eurl-ecvam-recommendation-on-the-direct-peptide-reactivity-assay-dpra.

(4) EC EURL ECVAM (2012), Direct Peptide Reactivity Assay (DPRA) Validation Study Report 74pp. Accessible at: http://ihcp.jrc.ec.europa.eu/our_labs/eurl-ecvam/eurl-ecvam-recommendations/eurl-ecvam-recommendation-on-the-direct-peptide-reactivity-assay-dpra

(5) Natsch *et al.* (2013), A dataset on 145 chemicals tested in alternative assays for skin sensitization undergoing prevalidation. Journal of Applied Toxicology, published online, 9 April 2013, DOI:10.1002/jat.2868.

(6) Jaworska *et al.* (2013), Bayesian integrated testing strategy to assess skin sensitization potency: from theory to practice Journal of Applied Toxicology, published online, 14 May 2013, DOI: 10.1002/jat.2869.

(7) DB-ALM (INVITTOX) Protocol 154: Direct Peptide Reactivity assay (DPRA) for skin sensitisation testing 17pp. Accessible at: http://ecvam-dbalm.jrc.ec.europa.eu/

(8) ECETOC (2003), Contact sensitization: Classification according to potency. European Centre for Ecotoxicology & Toxicology of Chemicals (Technical Report No. 87).

(9) Urbisch et al. (2016), Assessment of Pre- and Pro-haptens Using Nonanimal Test Methods for Skin Sensitization, Chem Res Toxicol. 29(5):901-13 doi: 10.1021/acs.chemrestox.6b00055

(10) Pattlewicz et al. (2016), Can currently available non-animal methods detect pre and pro-haptens relevant for skin sensitization? Regul Toxicol Pharmacol.; 82:147-155. doi: 10.1016/j.yrtph.2016.08.007

(11) Jaworska et al. (2015), Arch Toxicol. 2015 Dec;89(12):2355-83. doi: 10.1007/s00204-015-1634-2)

(12) OECD (2016). Series on Testing & Assessment No. 256: Guidance Document On The Reporting Of Defined Approaches And Individual Information Sources To Be Used Within Integrated Approaches To Testing And Assessment (IATA) For Skin Sensitisation, Annex 1 and Annex 2. ENV/JM/HA(2016)29. Organisation for Economic Cooperation and Development, Paris. Available at: [https://community.oecd.org/community/iatass].

APPENDIX I, ANNEX 1

PROFICIENCY SUBSTANCES

In Chemico Skin Sensitisation: Direct Peptide Reactivity Assay

Prior to routine use of the test method described in this test method, laboratories should demonstrate technical proficiency by correctly obtaining the expected DPRA prediction for the 10 proficiency substances recommended in Table 1 and by obtaining cysteine and lysine depletion values that fall within the respective reference range for 8 out of the 10 proficiency substances for each peptide. These proficiency substances were selected to represent the range of responses for skin sensitisation hazards. Other selection criteria were that they are commercially available, that high quality *in vivo* reference data and high quality *in vitro* data generated with the DPRA are available, and that they were used in the EURL ECVAM-coordinated validation study to demonstrate successful implementation of the test method in the laboratories participating in the study.

Table 1: Recommended proficiency substances for demonstrating technical proficiency with the Direct Peptide Reactivity Assay

Proficiency substances	CASRN	Physical state	*In vivo* prediction[1]	DPRA prediction[2]	Range[3] of % cysteine peptide depletion	Range[3] of % lysine peptide depletion
2,4-Dinitrochlorobenzene	97-00-7	Solid	Sensitiser (extreme)	Positive	90-100	15-45
Oxazolone	15646-46-5	Solid	Sensitiser (extreme)	Positive	60-80	10-55
Formaldehyde	50-00-0	Liquid	Sensitiser (strong)	Positive	30-60	≤24
Benzylideneacetone	122-57-6	Solid	Sensitiser (moderate)	Positive	80-100	≤7
Farnesal	19317-11-4	Liquid	Sensitiser (weak)	Positive	15-55	≤25
2,3-Butanedione	431-03-8	Liquid	Sensitiser (weak)	Positive	60-100	10-45
1-Butanol	71-36-3	Liquid	Non-sensitiser	Negative	≤7	≤5.5
6-Methylcoumarin	92-48-8	Solid	Non-sensitiser	Negative	≤7	≤5.5
Lactic Acid	50-21-5	Liquid	Non-sensitiser	Negative	≤7	≤5.5
4-Methoxyacetophenone	100-06-1	Solid	Non-sensitiser	Negative	≤7	≤5.5

[1]The *in vivo* hazard and (potency) predictions are based on LLNA data (5). The *in vivo* potency is derived using the criteria proposed by ECETOC (8).

[2] A DPRA prediction should be considered in the framework of an IATA and in accordance with the provisions of paragraphs 2 and 4.

[3] Ranges determined on the basis of at least 10 depletion values generated by 6 independent laboratories.

APPENDIX I, ANNEX 2

EXAMPLES OF ANALYSIS SEQUENCE

Calibration standards and reference controls	STD1 STD2 STD3 STD4 STD5 STD6 Dilution buffer Reference control A, rep 1 Reference control A, rep 2 Reference control A, rep 3
Co-elution controls	Co-elution control 1 for test chemical 1 Co-elution control 2 for test chemical 2
Reference controls	Reference control B, rep 1 Reference control B, rep 2 Reference control B, rep 3
First set of replicates	Reference control C, rep 1 Cinnamic aldehyde, rep 1 Sample 1, rep 1 Sample 2, rep 1
Second set of replicates	Reference control C, rep 2 Cinnamic aldehyde, rep 2 Sample 1, rep 2 Sample 2, rep 2
Third set of replicates	Reference control C, rep 3 Cinnamic aldehyde, rep 3 Sample 1, rep 3 Sample 2, rep 3
Reference controls	Reference control B, rep 4 Reference control B, rep 5 Reference control B, rep 6

Three sets of reference controls (i.e. samples constituted only by the peptide dissolved in the appropriate solvent) should be included in the analysis sequence:
Reference control A: used to verify the suitability of the HPLC system.
Reference control B: included at the beginning and at the end of the analysis sequence to verify stability of reference controls over the analysis time.
Reference control C: included in the analysis sequence to verify that the solvent used to dissolve the test chemical does not impact the percent peptide depletion.

APPENDIX II

In Chemico Skin Sensitisation: Amino acid Derivative Reactivity Assay (ADRA)

INITIAL CONSIDERATIONS, APPLICABILITY AND LIMITATIONS

1. The ADRA is proposed to address the molecular initiating event of the skin sensitisation AOP—namely, protein reactivity—by quantifying the reactivity of test chemicals towards model synthetic amino acid derivatives containing either lysine or cysteine (1) (2) (3). Depletion values of cysteine and lysine derivatives are then used to support the discrimination between skin sensitisers and non-sensitisers (1) (2) (3).

2. The ADRA proved to be transferable to laboratories experienced in high-performance liquid chromatography (HPLC) analysis. ADRA's WLR was 100% (10/10), 100% (7/7), 90% (9/10), and 100% (10/10) in four participating laboratories. BLR for 40 test chemicals calculated based the results from three participating laboratories was 91.9% (4). For the 40 chemicals tested in the validation study in four laboratories, the cumulative accuracy was 86.9% (139/160), sensitivity was 81.5% (88/108), and specificity was 98.1% (51/52) (4) (5). Results from the validation study (4) (5) as well as from other published studies (3) indicate that ADRA identified sensitisers and non-sensitisers with an accuracy of 79% (98/124) (124 compounds that fall within ADRA's applicability domain), a sensitivity of 74% (65/88), and a specificity of 92% (33/36) relative to LLNA results (6). In addition, the prediction of human skin sensitisation for 73 compounds that fall within ADRA's applicability domain has an accuracy of 86% (63/73), a sensitivity of 85% (44/52), and a specificity of 90% (19/21) (6). However, the accuracy values given here for ADRA as a stand-alone test method are for reference only, since it is recommended that the test method be used in combination with other sources of information in the context of an IATA and in accordance with the provisions of paragraphs 7 and 8 in the General Introduction. Furthermore when evaluating non-animal methods for skin sensitisation, it should be kept in mind that the LLNA test as well as other animal tests may not fully reflect the situation in the species of interest, which is humans. On the basis of the overall data available, ADRA's applicability domain was shown to include a variety of organic functional groups, reaction mechanisms, skin sensitisation potencies (as determined in *in vivo* studies), and physicochemical properties (1) (2) (3) (4). Following an independent peer review, the ADRA validation study was considered to demonstrate that this method should be acceptable as part of an integrated testing strategy for the predictive identification of skin sensitisation hazard (7).

3. The term "test chemical" is used in this Test Guideline to refer to what is being tested and is not related to the applicability of the ADRA to the testing of substances and/or mixtures. This test method is not applicable to the testing of metal compounds, which are known to react with proteins via mechanisms other than covalent binding. The test method described in this Appendix of the Test Guideline is an *in chemico* method that does not encompass a metabolic system. Chemicals that require enzymatic bioactivation to exert their skin sensitisation potential (i.e. pro-haptens) cannot be detected by the test method. Chemicals that become sensitisers after abiotic transformation (i.e. pre-haptens) are

reported to be in some cases correctly detected by the test method (1) (2) (3) (4). In the light of the above, negative results obtained with the test method should be interpreted in the context of the stated limitations and in the connection with other information sources within the framework of an IATA. Test chemicals that promote the oxidation of the *N*-(2-(1-naphthyl)acetyl)-L-cysteine (NAC) reagent (i.e. cysteine dimerisation) could lead to a potential over-estimation of NAC depletion, resulting in possible false positive predictions (see paragraphs 27 and 28); it may be possible to detect and quantify any NAC dimer formed by HPLC, thus confirming or ruling out that the NAC reagent has been depleted via oxidative dimerisation as opposed to reaction and covalent bonding to the test item substance(s).

4. The ADRA test method allows testing of poorly soluble chemicals. To be tested, a test chemical should be soluble in an appropriate solvent at a final concentration of 1 mM (see paragraph 14). Test chemicals that are not soluble at this concentration may still be tested at lower concentrations. In such cases, a positive result could still be used to support identification of the test chemical as a skin sensitiser but no firm conclusion on the lack of reactivity should be drawn from a negative result.

5. In general, many organic compounds absorb UV in the range of 220 nm. In the case of co-elution of the nucleophilic reagent and the test chemical, this might result in false negative prediction. This may happen with the DPRA which specifies that quantification of the peptide-based nucleophilic reagents has to be performed at 220 nm. In contrast to this, the nucleophilic reagents used in ADRA are quantified at 281 nm. The substances that absorb UV in this range of the spectrum are generally limited to those having conjugated double bonds, which significantly lowers the potential for co-elution (8).

6. The current prediction model cannot be used for complex mixtures of unknown composition or for substances of unknown or variable composition, complex reaction products, or biological materials (UVCB substances) due to the need for defined molar ratio of test chemical and nucleophilic reagents. Limited information is currently available on the applicability of the ADRA to mixtures (9) (10). A new protocol has to be developed for multi-constituent substances and mixtures to be used with test methods like ADRA, which utilise HPLC analysis to quantify the depletion of nucleophilic reagents (9) (10). Thus, although it is impossible to define fixed methods in this guideline, which can evaluate multi-constituent substances and mixtures, paragraph 15 describes an evaluation method that is considered to be applicable at the present time for multi-constituent substances or mixtures of known composition (9). Such substances were however not tested during the validation studies. When considering testing of mixtures, difficult-to-test chemicals (e.g. unstable), or test chemicals not clearly within the applicability domain described in this Guideline, upfront consideration should be given to whether the results of such testing will yield results that are meaningful scientifically.

7. ADRA can be used to support the discrimination between skin sensitisers and non-sensitisers. Further work, preferably based on human data, is necessary to determine whether ADRA results can contribute to potency assessment when considered in combination with other information sources.

PRINCIPLE OF THE TEST

8. ADRA is an *in chemico* test method that quantifies residual concentrations of the cysteine derivative *N*-(2-(1-naphthyl)acetyl)-*L*-cysteine (CAS. 32668-00-1), which is known as NAC, and the lysine derivative α-*N*-(2-(1-naphthyl)acetyl)-*L*-lysine (CAS. 397841-92-8), known as NAL, following a 24±1 hour incubation at 25±1ºC in the presence of a test chemical. Both these derivatives include a naphthalene ring that is introduced to their *N*-terminal in order to facilitate UV detection. The relative concentrations of NAC and NAL are measured by high-performance liquid chromatography (HPLC) with gradient elution and UV detection at 281 nm. Percent depletion values are then calculated for both NAC and NAL and compared to a prediction model (see paragraph 26).

9. Prior to routine use of the method described in this test method, laboratories should demonstrate technical proficiency, using the ten proficiency substances listed in Annex 1 of this Appendix.

PROCEDURE

10. This test method is based on the protocol (11) used for the JaCVAM-coordinated ADRA validation study and is recommended for use when implementing ADRA at a laboratory. The main components and procedures for the ADRA are described below. Before using an alternative HPLC set-up, its equivalence to the validated set-up described in the protocol should be demonstrated, preferably by testing the proficiency substances in Annex 1 of this Appendix.

Quality of NAC and NAL

11. The Nucleophilic Reagents can be obtained as an ADRA Kit for Skin Sensitisation Test, from FUJIFILM Wako (FFWK) Pure Chemical Corporation, Catalog No. 296-80901. Manufacturing NAC/NAL is patented in Japan only, by Fujifilm Corporation. Therefore, manufacturers in other countries can produce NAC/NAL without permission. In case other NAC/NAL are used, these should satisfy three quality criteria described below. Quality checks can be obviated and ADRA testing performed without delay by purchasing NAC and NAL that have been manufactured specifically to satisfy these quality criteria.

Quality required for NAC and NAL:

1) Purity: Both NAC and NAL are to be at least 98% pure.

2) Stability: Using NAC and NAL stock solution, prepare a reference control free of any test chemical and quantify the residual levels of NAC and NAL both immediately after preparation (0 hours) and after a 24 hour incubation. Residual levels of NAC and NAL are to be a minimum of 90% in either case (11). The residual level of NAC is calculated as a percentage of the sum of NAC and the residual level of NAC dimers.

3) Reactivity: NAC and NAL are to be evaluated for reactivity with the ten proficiency substances given in Annex 1 and should satisfy the requirement given therein.

Preparation of the NAC and NAL stock solution

12. The solubility of individual NAC and NAL batches should be verified prior to use. NAC stock solution should be prepared to a concentration of 2 mM in 100 mM of pH 8.0 phosphate buffer, including 0.333 μM of EDTA, as well as NAL stock solution to a concentration of 2 mM in 100 mM of pH 10.2 phosphate buffer. These two stock solutions are then diluted in buffer to prepare 6.667 μM stock solutions. Both NAC and NAL stock solutions should be used as soon as possible after preparation (3). In the event that they are to be stored, these stock solutions may be frozen and stored for up to twelve months time at less than -75°C prior to use. The final concentration of the NAC solution is 5 μM in pH 8.0 phosphate buffer, and the final concentration of the NAL solution is 5 μM in pH 10.2 phosphate buffer.

Preparation of the test chemical solution

13. Solubility of the test chemical in an appropriate solvent should be assessed before performing the assay in accordance with the solubilisation procedure described in the ADRA JaCVAM protocol (11). An appropriate solvent should dissolve the test chemical completely. Since the ADRA protocol stipulates that the test chemical be incubated in an excess volume of both NAC and NAL, visual inspection of the clear test chemical solution is considered sufficient to confirm that the test chemical (and all its constituents, if testing a multi-constituent substance or a mixture) is dissolved. Suitable solvents are distilled water, acetonitrile and acetone. If the test chemical is not soluble in any of the solvents mentioned above, DMSO can be used as a last resort and in minimal amounts. It is important to note that DMSO may lead to dimerisation of the nucleophilic reagent NAC (12) and as a result, it may be more difficult to meet the acceptance criteria. If DMSO is chosen, attempts should be made to solubilise the test chemical in a 1:20 mixture of DMSO and acetonitrile (5% DMSO in acetonitrile). When using a DMSO-acetonitrile solvent, the test chemical should be dissolved in DMSO, and then this solution should be diluted 20-fold with acetonitrile to prepare a 1 mM test chemical solution. In case the use of DMSO leads to increased dimerisation of the NAC reagent, this can be checked analytically as the NAC dimer can be detected by HPLC. The test chemical should be pre-weighed into a disposable polypropylene tube and dissolved immediately before testing in an appropriate solvent to prepare a 1 mM solution.

14. Mono-constituent substances of unknown molecular weight may be tested in a test chemical solution at a concentration of 0.5 mg/mL rather than 1 mM (9). Polymers which are well characterised should also be tested at a concentration of 1 mM based on the mean number average molecular weight, in a manner analogous to the procedure for mono-constituent compounds.

15. Mixtures and multi constituent substances, of known composition are to be tested as follows:

> 1) Liquids: Generally, tested as an undiluted mixture. In cases where low solubility of the test item prevents formation of reaction solution, i.e. undissolved material, clouding, and/or precipitation is observed, a positive result may still be used in the assessment, whereas a negative result is uncertain and should be interpreted with due care. Insofar as results could be false positives, however, predictions should be interpreted with due care.
>
> 2) Solids: The test chemical should be dissolved to maximum soluble concentration in the same solvent used to prepare the 1 mM test chemical solution. The test

chemical solution of the highest concentration possible is then tested as an undiluted mixture. In cases where low solubility of the test item prevents formation of reaction solution, i.e. undissolved material, clouding, and/or precipitation is observed, a positive result may still be used in the assessment, whereas a negative result is uncertain and should be interpreted with due care. Insofar as results could be false positives, however, predictions should be interpreted with due care.

Preparation of the positive control, reference controls and co-elution controls

16. Phenylacetaldehyde (CAS 122-78-1, ≥90% purity) should be used as positive control (PC) at a concentration of 1 mM in acetonitrile. Other suitable positive controls that provide mid-range depletion values may be used if historical data are available to derive comparable run acceptance criteria. In addition, reference controls comprising only NAC or only NAL dissolved in the appropriate solvent should also be included in the HPLC run sequence, so they can be used to verify the HPLC system suitability prior to analysis (Reference Control A), the stability of the reference controls over time (Reference Control B), and any effects of the solvent used on depletion of NAC or NAL (Reference Control C) (See Annex 2). The percent NAC and NAL depletion for a test chemical is calculated using an appropriate reference control for that test chemical (see paragraph 23). Also, a co-elution control comprising only the test chemical should be included in the run sequence to detect possible co-elution of the test chemical with either the NAC or NAL.

Incubation of the test chemical with the NAC and NAL solutions

17. Both the NAC and the NAL solutions should be incubated with the test chemical at 1:50 ratio in a 96-well microplate. The observation of precipitate immediately upon addition of the test chemical solution to the NAC and the NAL solutions is an indication of poor solubility, which means that there is no way to know exactly how much test chemical is contained in the solution. Thus, although positive results can be used with confidence, negative results are uncertain and should be interpreted with due care (see also paragraph 4 regarding the testing of chemicals not soluble at concentrations as high as 1 mM). The reaction solution should be incubated in the dark at 25±1ºC for 24±1 hours before performing HPLC analysis. After incubation, trifluoroacetic acid (TFA) (≥ 98%) should be added as a fixing solution to stop the reaction (3).

HPLC preparation and analysis

18. Each test chemical should be analysed in triplicate to determine percent depletion for both NAC and NAL. Although adding the fixing solution does stop the reaction, measurement of the reaction solution is to be performed as soon as possible and in any case within three days after adding the fixing solution. For example, when HPLC analysis of NAC and NAL are performed separately using two 96-well microplates, up to 34 samples may be analysed at one time, including the test chemical, the positive control, and the appropriate number of solvent controls based on the number of individual solvents used in the test, each in triplicate. All of the replicates analysed in a single run should use identical batches of NAC and NAL stock solution. Test chemical and control solutions are to be visually inspected prior to HPLC analysis and may be centrifuged at low speed (100–400 × g) to force any precipitate to the bottom of the vial as a precaution against large amounts of precipitate clogging the HPLC tubing or columns. Observation of precipitation or phase separation after the incubation period is an indication that NAC and NAL depletion could be misleading, and negative results in that case are uncertain and should be interpreted with

due care, as well as for any precipitate observed at the beginning of the incubation period (see above).

19. A standard calibration curve should be generated for both NAC and NAL. Standard solutions of both NAC and NAL should be prepared in 20% acetonitrile in buffer and containing 0.5% trifluoroacetic acid. For NAC, a phosphate buffer at pH 8.0, and for NAL, a phosphate buffer at pH 10.2 should be used. Serial dilution of the NAC and NAL stock solutions (5.0 μM) will be used to prepare six calibration solutions in concentrations from 5.0 to 0.156 μM as well as a blank of the dilution buffer. Suitable calibration curves should have an $R^2 > 0.990$.

20. The suitability of the HPLC system should be verified before conducting the analysis. Both NAC and NAL depletion is monitored by HPLC coupled with an UV detector (photodiode array detector or fixed wavelength absorbance detector with 281 nm signal). The appropriate column is installed in the HPLC system. The recommended HPLC set-up described in the validated protocol uses a column (Base particle: core-shell type silica gel, Particle size: 2.5~2.7 μm, colomn size: 3.0 × 150 mm) as preferred column. With this reversed-phase HPLC column, the entire system should be equilibrated for at least 30 minutes at 40ºC with 50% phase A (0.1% (v/v) trifluoroacetic acid in water), 50% phase B (0.1% (v/v) trifluoroacetic acid in acetonitrile) before use. Then, the column is conditioned by running the gradient at least twice before actual use. The HPLC analysis should be performed using a flow rate of 0.30 mL/min and a linear gradient from 30% to 55% acetonitrile for NAC and from 25% to 45% acetonitrile for NAL within 10 minutes, followed by a rapid increase to 100% acetonitrile to remove other materials. Equal volumes of the standard solutions, test chemical solutions, and control solutions should be injected. The column should be re-equilibrated under initial conditions for 6.5 minutes between injections. If a different reversed-phase HPLC column is used, the set-up parameters described above may need to be adjusted to guarantee an appropriate elution and integration of the NAC and NAL, including the injection volume, which may vary according to the system used (typically in the range from 10–20 μL). Importantly, if an alternative HPLC set-up is used, its equivalence to the validated set-up described above should be demonstrated, preferably by testing the proficiency substances in Annex 1. Absorbance is monitored at 281 nm. If a photodiode array detector is used, absorbance at 291 nm should also be recorded. It should be noted that some batches of acetonitrile could have a negative impact on NAC and NAL stability and this has to be assessed when a new batch of acetonitrile is used. The ratio of the 281 nm peak area and the 291 nm peak area can be used as an indicator of co-elution. For each sample a ratio in the range of 90% < mean area ratio of control samples < 100% would give a good indication that co-elution has not occurred. An example of HPLC analysis sequence is provided in Annex 2.

21. There are some test chemicals that could potentially promote oxidation of NAC. The peak of the dimerised NAC may be monitored visually. Any apparent dimerisation should be noted, since overestimation of NAC depletion could result in false-positive predictions (See paragraphs 26 and 27).

DATA AND REPORTING

Data evaluation

22. The concentration of both NAC and NAL is photometrically determined at 281 nm in each sample by measuring the peak area (area under the curve, AUC) of the appropriate peaks and by calculating the concentration of both NAC and NAL using the linear calibration curve derived from the standards.

23. The percent depletion for both NAC and NAL is determined in each sample by measuring the peak area and dividing it by the mean peak area of the relevant Reference Controls C (See Annex 2) according to the formula described below.

$$\text{Percent NAC or NAL depletion} = \left[1-\left[\frac{\text{NAC or NAL peak area in replicate injection}}{\text{Mean NAC or NAL peak area in reference controls C}}\right]\right]\text{x}100$$

Acceptance criteria

24. The following criteria should be met:

a) the standard calibration curve should have an $R^2 > 0.990$,

b) the mean percent NAC and NAL depletion value of the three replicates for the positive control phenylacetaldehyde should be between 6% and 30% for NAC and between 75% and 100% for NAL, while the maximum standard deviation (SD) for the positive control replicates should be < 10% for both NAC and NAL depletion, and

c) the mean NAC and NAL concentration of both Reference Controls A and Reference Control C should be 3.2–4.4 µM and the coefficient of variation (CV) of NAC and NAL peak areas for the nine Reference Controls B and C in acetonitrile should be < 10%.

If one or more of these criteria is not satisfied, the data should be rejected and the run should be repeated.

25. The following criteria should be satisfied for a test chemical's results to be accepted as valid:

a) the maximum standard deviation for the test chemical replicates should be < 10% for the percent depletion of both NAC and NAL,

b) the mean NAC and NAL concentration of the three Reference Controls C in the appropriate solvent should be 3.2–4.4 µM.

If one or more of these criteria is not satisfied, the data should be rejected and the run should be repeated.

Prediction model

26. The mean percent depletion of NAC and NAL is calculated for each test chemical. Negative depletion is considered to be "0" when calculating the mean. By using the NAC/NAL prediction model shown in Table 1, the threshold of 4.9% mean percent depletion should be used to support the discrimination between skin sensitisers and non-sensitiser in the framework of an IATA or a DA.

Table 1: NAC/NAL prediction model[1]

Mean NAC and NAL percent depletion	ADRA prediction[2]
Less than 4.9%	Negative
4.9% or higher	Positive

[1] The numbers refer to statistically generated threshold values and are not related to the precision of the measurement.
[2] An ADRA prediction should be considered in the framework of an IATA and in accordance with the provisions of paragraphs 2 and 3.

27. Co-elution occurs when the test chemical (the substance or one or several of the constituents of a multi-constituent substance or a mixture) absorbs significantly at 281 nm and has the same retention time as NAC or NAL. Co-elution may be resolved by slightly adjusting the HPLC set-up in order to further separate the elution time of the test chemical and NAC or NAL. If an alternative HPLC set-up is used to try to resolve co-elution, its equivalence to the validated set-up should be demonstrated, preferably by testing the proficiency substances in Annex 1. When co-elution occurs, it is not possible to integrate the peak of the NAC or NAL, thereby preventing calculation of the percent depletion of NAC or NAL. If co-elution of test chemicals occurs with both the NAC and NAL and separation of elution time is not feasible, then the analysis should be reported to be inconclusive. In cases where co-elution occurs only with NAL and separation of elution time is not feasible, the NAC-only prediction model (See Table 2) can be used to make a prediction.

Table 2: NAC-only prediction model[1]

Mean NAC percent depletion	ADRA prediction[2]
Less than 5.6%	Negative
5.6% or higher	Positive

[1] The numbers refer to statistically generated threshold values and are not related to the precision of the measurement.
[2] An ADRA prediction should be considered in the framework of an IATA and in accordance with the provisions of paragraphs 2 and 3.

28. When a result is unequivocal, a single HPLC analysis for both NAC and NAL should be sufficient for a test chemical. Additional testing is sometimes necessary, however, when the results lie close to the threshold value used to discriminate between positive and negative results (borderline results). If the mean percent depletion falls between 3.0% and 10.0% when using the NAC/NAL prediction model or the NAC percent depletion falls between 4.0% and 11.0% when using the NAC-only prediction model, a second run is advisable, as is a third run in the event of discordant results between the first two runs.

Test report

29. The test report should include the following information:

Test chemical

- Mono-constituent substance
    - Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers
    - Physical appearance, water solubility, molecular weight, and additional relevant physicochemical properties, to the extent available
    - Purity, chemical identity of impurities as appropriate and practically feasible, etc.
    - Treatment prior to testing, if applicable (warming, grinding)
    - Concentration(s) tested
    - Storage conditions and stability to the extent available
- Multi-constituent substance, UVCB, and mixtures
    - Characterisation by chemical identity (see above), purity, quantitative occurrence and relevant physicochemical properties (see above) of the constituents, to the extent available
    - Physical appearance, water solubility, and additional relevant physicochemical properties, to the extent available
    - Molecular weight (or apparent molecular weight) for mixtures or polymers of known composition, or other information relevant to the study
    - Treatment prior to testing, if applicable (warming, grinding)
    - Concentration(s) tested
    - Storage conditions and stability, to the extent available.

Controls

- Positive control
    - Chemical identification, such as IUPAC or CAS name(s), CAS number(s), SMILES or InChI code, structural formula, and/or other identifiers;
    - Physical appearance, water solubility, molecular weight, and additional relevant physicochemical properties, to the extent available;
    - Purity, chemical identity of impurities, as appropriate or feasible
    - Treatment prior to testing, if applicable (warming, grinding)
    - Concentration(s) tested
    - Storage conditions and stability, to the extent available;
    - Reference to historical positive control results demonstrating suitable run acceptance criteria, if applicable.

- Solvent
    - Solvent used and ratio of its constituents, if applicable
    - Chemical identification(s), such as IUPAC or CAS name(s), CAS number(s), and/or other identifiers
    - Purity, chemical identity of impurities, as appropriate and feasible
    - Physical appearance, molecular weight, and additional relevant physicochemical properties when solvents other than those mentioned in the test method are used
    - Storage conditions and stability, to the extent available
    - Justification for choice of solvent for each test chemical
    - Impact on NAC and NAL stability when using acetonitrile

Preparation of NAC and NAL, positive control and test chemical solution

- Characterisation of NAC and NAL solutions (supplier, lot, exact weight of NAC and NAL, volume added for the stock solution)
- Characterisation of positive control solutions (exact weight of positive control reagent, volume added for the control solution)
- Characterisation of test chemical solutions (exact weight of test chemical, volume added for the test chemical solution)

HPLC instrument setting and analysis

- Type of HPLC instrument, HPLC and guard columns, detector, autosampler
- Parameters relevant for the HPLC analysis such as column temperature, injection volumes, flow rate and gradient

System suitability

- NAC and NAL peak area at 281 nm of each standard and reference control A replicate
- Linear calibration curve graphically represented and the R2 reported
- NAC and NAL concentration of each Reference Control A replicate
- Mean NAC and NAL concentration (μM) of the three reference controls A, SD and CV
- NAC and NAL concentration of Reference Controls A and C.

Analysis sequence

- For Reference Controls
    - NAC and NAL peak area at 281 nm of each replicate of Reference Controls B and C
    - Mean NAC and NAL peak area at 281 nm of the nine Reference Controls B and C in acetonitrile, SD and CV (for stability of reference controls over analysis time)

- For each solvent used, the mean NAC and NAL peak area at 281 nm of the three appropriate Reference Controls C (for the calculation of percent NAC and NAL depletion)
- For each solvent used, the NAC and NAL concentration (μM) of the three appropriate Reference Controls C
- For each solvent used, the mean NAC and NAL concentration (μM) of the three appropriate Reference Controls C, SD and CV.

- For positive controls
  - NAC and NAL peak area at 281 nm of each replicate
  - Percent NAC and NAL depletion of each replicate
  - Mean percent NAC and NAL depletion of the three replicates, SD and CV.

- For each test chemical
  - Appearance of precipitate in the reaction mixture at the end of the incubation time, if observed. If precipitate was re-solubilised or centrifuged;
  - Presence of co-elution
  - Description of any other relevant observations, if applicable
  - NAC and NAL peak area at 281 nm of each replicate
  - Percent NAC and NAL depletion of each replicate
  - Mean of percent NAC and NAL depletion of the three replicate, SD and CV
  - Mean of percent NAC and percent NAL depletion values
  - Prediction model used and ADRA prediction

Proficiency testing

- If applicable, the procedure used to demonstrate proficiency of the laboratory in performing the test method (testing of proficiency substances, etc.) or to demonstrate reproducible performance of the test method over time.

Discussion of the results

- Discussion of the results obtained with the ADRA test method
- Discussion of the test method results in the context of an IATA if other relevant information is available

Conclusion

Literature for Appendix II

(1) Fujita M, Yamamoto Y, Tahara H, Kasahara T, Jimbo Y and Hioki T (2014), Development of a prediction method for skin sensitisation using novel cysteine and lysine derivatives, Journal of Pharmacological and Toxicological Methods, 70:94-105. DOI: 10.1016/j.vascn.2014.06.001.

(2) Yamamoto Y, Tahara H, Usami R, Kasahara T, Jimbo Y, Hioki T and Fujita M (2015), A novel *in chemico* method to detect skin sensitisers in highly diluted reaction conditions, Journal of Applied Toxicology, 35:1348-60. DOI: 10.1002/jat.3139.

(3) Fujita M, Yamamoto Y, Watanabe S, Sugawara T, Wakabayashi K, Tahara K, Horie N, Fujimoto K, Kusakari K, Kurokawa Y, Kawakami T, Kojima K, Kojima H, Ono A, Katsuoka Y, Tanabe H, Yokoyama H and Kasahara T (2019), Cause of and Countermeasures for Oxidation of the Cysteine-Derived Reagent Used in the Amino acid Derivative Reactivity Assay, Journal of Applied Toxicology, 39,191-208. DOI: 10.1002/jat.3707.

(4) Fujita M, Yamamoto Y, Watanabe S, Sugawara T, Wakabayashi K, Tahara Y, Horie N, Fujimoto K, Kusakari K, Kurokawa Y, Kawakami T, Kojima K, Sozu T, Nakayama T, Kusao T, Jon R, Kleinstreuer N, Bae-Hwa K, Kojima H, Kasahara T, Ono A, The within- and between-laboratory reproducibility and predictive capacity of the in chemico Amino acid Derivative Reactivity Assay (ADRA): Results of validation study implemented in four participating laboratories, Journal of Applied Toxicology, accepted on May 10, 2019

(5) OECD (2019), Draft validation report: Amino acid Derivative Reactivity Assay (ADRA) – JaCVAM Validation Study Report. Series on testing and Assessment n° 304. Organisation for Economic Cooperation and Development, Paris.

(6) Yamamoto Y, Fujita M, Wanibuchi S, Sato A, Katsuoka Y, Ono A, Kasahara T. (2019), Availability of ADRA for prediction of skin sensitization by combining multiple alternative methods. Journal of Toxicological Sciences, accepted on May 9, 2019

(7) OECD (2019), Amino acid Derivative Reactivity Assay (ADRA) – Report of the Peer Review Panel. Series on testing and Assessment n° 305. Organisation for Economic Cooperation and Development, Paris.

(8) Fujira M, Yamamoto Y, Wanibuchi S, Katsuoka Y, Kasahara T, (2019), The underlying factors that explain why nucleophilic reagents rarely co-elute with test chemicals in the ADRA. Journal of Pharmacological and Toxicological Methods. 96, 95-105. DOI: 10.1016/j.vascn.2019.02.004

(9) YamamotoY, Fujita M, Manibuchi S, Katsuoka Y, Ono A, Kasahara T.. (2019), Expanding the applicability of the Amino acid Derivative Reactivity Assay (ADRA): Determining a weight concentration for preparation of test chemical solutions that yields a predictive capacity identical to the conventional method using molar concentration and demonstrating the capacity to detect sensitizers in liquid mixtures. Journal of Pharmacological and Toxicological Methods. 97, 67-79. DOI: 10.1016/j.vascn.2019.01.001

(10) Fujita M, Yamamoto Y, Wanibuchi S, Katsuoka Y, Kasahara T A Newly Developed Means of HPLC-Fluorescence Analysis for Predicting the Skin Sensitization Potential of

Multi-Constituent Substances Using ADRA. Toxicology In Vitro, 59, 161-178. DOI: 10.1016/j.tiv.2019.04.014.

(11) ADRA protocol: JaCVAM Statements. Available at: http://www.jacvam.jp/en_effort/effort02.html

(12) James P. Tam JP, Cui Rong Wu CR, Wen Liu W, Jing Wen Zhang JW. (1991), Disulfide bond formation in peptides by dimethyl sulfoxide. Scope and applications. Journal of the American Chemical Society, 113, 6657–6662. DOI: 10.1021/ja00017a044

(13) Natsch A, Ryan CA, Foertsch L, Emter R, Jaworska J, Gerberick F, Kern P. (2013). A dataset on 145 chemicals tested in alternative assays for skin sensitisation undergoing prevalidation. *Journal of Applied Toxicology*, **33**(11):1337-52, DOI:10.1002/jat.2868.

(14) Gerberick GF, Vassalo JD, Foertsch LM, Price BB, Chaney JG, Lepoittevin J-P, (2007), Quantification of chemical peptide reactivity for screening contact allergens: A classification tree model approach, Toxicological Sciences, 97, 417-427. DOI: 10.1093/toxsci/kfm064.

(15) Basketter DA, Scholes EW, (1992), Comparison of the local lymph node assay with the guinea-pig maximization test for the detection of a range of contact allergens, Food and Chemical Toxicology, 30, 65-69.

(16) ECETOC (2003). Contact sensitisation: Classification according to potency. European Centre for Ecotoxicology & Toxicology of Chemicals (Technical Report No. 87).

APPENDIX II, ANNEX 1

Proficiency Substances

In Chemico Skin Sensitisation: Amino acid Derivative Reactivity Assay (ADRA)

Prior to routine use of the test method, laboratories should demonstrate technical proficiency by correctly obtaining the expected ADRA prediction for the 10 proficiency substances recommended in Table 1 and by obtaining NAC and NAL depletion values that fall within the respective reference ranges for 8 out of the 10 proficiency substances. These proficiency substances were selected to represent the full range of responses for skin sensitisation hazards. Other selection criteria were that they are commercially available, that high quality *in vivo* reference data and high quality ADRA data are available, and that they were used during the JaCVAM-coordinated validation study to demonstrate successful implementation.

Table 1. Recommended chemicals for demonstrating technical proficiency with ADRA

No.	Test chemicals	CAS No.	Physical state	Molecular weight	*In vivo* Prediction[1]	ADRA prediction[2]	Range of % depletion	
							NAC[3]	NAL[3]
1	*p*-Benzoquinone	106-51-4	Solid	108.09	Sensitiser (extreme)	Positive	90-100	40-70
2	Chloramine T trihydrate	7080-50-4	Solid	281.69	Sensitiser (strong)	Positive	90-100	90-100
3	Trans-Cinnamaldehyde	14371-10-9	Liquid	132.16	Sensitiser (moderate)	Positive	40-100	≤20
4	Palmitoyl Chloride	112-67-4	Liquid	274.87	Sensitiser (moderate)	Positive	≤10	50-100
5	Imidazolidinyl urea	39236-46-9	Solid	388.29	Sensitiser (weak)	Positive	10-45	≤10
6	Farnesal	19317-11-4	Liquid	220.35	Sensitiser (weak)	Positive	20-40	≤15
7	Glycerol	56-81-5	Liquid	92.09	Non-sensitiser	Negative	≤7	≤7
8	Benzyl alcohol	100-51-6	Liquid	108.14	Non-sensitiser	Negative	≤7	≤7
9	Dimethyl isophthalate	1459-93-4	Solid	194.19	Non-sensitiser	Negative	≤7	≤7
10	Propyl paraben	94-13-3	Solid	180.20	Non-sensitiser	Negative	≤7	≤7

[1]The *in vivo* hazard (and potency) predictions are based on LLNA data. (13) (14) (15). The *in vivo* potency is derived using the criteria proposed by ECETOC (16).

[2] An ADRA prediction should be considered in the framework of an IATA and in accordance with the provisions of paragraphs 2 and 3.

[3] Ranges determined on the basis of at least 10 depletion values generated by 5 independent laboratories.

APPENDIX II, ANNEX 2

EXAMPLES OF ANALYSIS SEQUENCE

Each sample of HPLC analysis should be analysed in number order below. Refer to the table showing Examples of HPLC Sample Analysis Sequences for more practical sequences about HPLC analysis.

1. Start to analyse calibration standards and Reference Control A (N = 3).

2. The co-elution Control does not need to be analysed by turns if it is analysed after analysis of standard solution and Reference Control A.

3. Reference Control B should be analysed three times (total six times) before and after the analysis of sample, Reference Control C and Positive Control.

4. The Reference Control C, Positive Control and Test chemical solutions are analysed. (After the first set of replicates of each sample is analysed, the second set of replicates of each should be analysed).

Calibration standards and reference controls	STD1 STD2 STD3 STD4 STD5 STD6 Dilution buffer Reference control A, rep 1 Reference control A, rep 2 Reference control A, rep 3
Co-elution controls	Co-elution control 1 for test chemical 1 Co-elution control 2 for test chemical 2
Reference controls	Reference control B, rep 1 Reference control B, rep 2 Reference control B, rep 3
First set of replicates	Reference control C, rep 1 Phenylacetaldehyde, rep 1 Sample 1, rep 1 Sample 2, rep 1
Second set of replicates	Reference control C, rep 2 Phenylacetaldehyde, rep 2 Sample 1, rep 2 Sample 2, rep 2
Third set of replicates	Reference control C, rep 3 Phenylacetaldehyde, rep 3 Sample 1, rep 3 Sample 2, rep 3
Reference controls	Reference control B, rep 4 Reference control B, rep 5 Reference control B, rep 6

Three sets of reference controls (NAC or NAL dissolved in the appropriate solvent) should be included in the analysis sequence:

Reference control A: Control for verifying validity of the HPLC system. Reference Control A is used to verify concentration of NAC and NAL from each calibration curve after addition of acetonitrile rather than test chemical.

Reference control B: Control for verifying stability of reaction solution under analysis. Reference Control B is used to verify variability (CV) of each three NAC/NAL peak areas in the solution after addition of acetonitrile rather than test chemical at the start of analysis and at the end of analysis.

Reference control C:

Control for calculating NAC/NAL depletion of each test chemical solution. To calculate depletion of NAC/NAL, measure three Reference Controls C after addition of solvent instead of test chemical. Prepare reference Control C for all solvents used to dissolve the test chemicals.

IL-2 Luciferase (IL-2 Luc) Assay Report of the Peer Review Panel

on

a JaCVAM co-ordinated study programme addressing the validation status of the IL-2 Luc assay for the prospective identification of immunotoxic substances on T cells

Report completed by the Peer review Panel on June 30th, 2020

Table of Contents

Executive Summary

The IL-2 luciferase (Luc) assay has been proposed as an in vitro alternative to animal testing. The assay provides information on adverse outcome pathways for immunotoxicity to T cells.
The Peer Review Panel (PRP) found the Validation Management Team's report presented the necessary information for an independent review.
The PRP concluded that the IL-2 Luc assay was well defined and has a clear protocol and criteria for data interpretation. All necessary information, including performance standards, was sufficiently detailed. Both within- and between-laboratory reproducibility were satisfactory. The PRP noted that the use of compounds with a better-defined immunotoxicity mechanism will help to improve the accuracy of the assay. While the predictive capacity was not satisfactory for a stand-alone method, the IL-2 Luc assay is acceptable for use in an Integrated Approach to Testing and Assessment.

Peer Review Panel Composition

Fujio Kayama (Chair)	Jichi Medical University, Japan
Henk van Loveren	Maastricht University, The Netherlands
Haley Neff-LaFord	Seattle Genetics, Inc., USA
Barbara Kaplan	Mississippi State University, USA
Xingchao Geng	National Center for Safety Evaluation of Drugs (NCSED), China
Takao Ashikaga	JaCVAM, Kawasaki, Japan
Sang-Hyun Kim	Kyungpook National University, Korea

Background

Immune dysregulation can have serious adverse health consequences and it could be caused by many types of chemicals, such as environmental contaminants, food additives, and drugs. It ranges from reduced resistance to infection and neoplasia to allergic and autoimmune conditions. For many years, to identify such immunotoxic chemicals depended on animal models. For animal welfare, ethical and scientific reasons there has been a desire to replace in vivo methods with non-animal alternatives (1). The immune system comprises innate and adaptive immunity. Both arms of the immune response function differently and are driven by different population of cells. A variety of intracellular signaling pathways also play roles in innate and adaptive immune responses. Given the complexity of the immune system, it is unlikely that a single in vitro method will cover all immunotoxicants. Therefore, Integrated Approaches to Testing and Assessment (IATA) are flexible approaches for safety assessment based on the integration and translation of the data derived from multiple methods and sources. For example, target compounds are tested using various approaches such as the human whole blood cytokine release assay (HWBCRA), lymphocyte proliferation assay, mixed lymphocyte reaction and fluorescent cell chip assay after the evaluation of myelotoxicity (2). There are many in vitro methods aiming at evaluating various aspects of molecular and cellular events in the adverse outcome pathway (AOP) for immunotoxicity. However, no in vitro method is validated formally.

The IL-2 Luc assay was developed to be part of a high-throughput screening system that enables evaluation of chemical immunotoxicity. This screening system was named Multi-ImmunoTox assay or MITA (3) (4) (5). The IL-2 Luc assay, using a human cell line transfected with luciferase genes under control of the IL-2 promoter, identifies the effect of chemicals on the IL-2 activity in the 2H4 cells in the presence of stimulants (3).

The PRP first met in February 2019 to review a progress report on the IL-2 Luc assay prepared by the Validation Management Team (VMT). Following commentary on this work by the PRP, the VMT refined the validation report.
The PRP engaged in follow-up telephone conferences in October 2019, November 2019 and June 2020. With the provision of all of the amended, updated and additional material, including the final VMT report, this PRP Report was prepared.

IL-2 Luc Test Method Definition

The PRP confirmed that the IL-2 Luc assay test method has been fully described in the

report of the VMT and in the associated detailed test protocol. During the validation study, the test developer changed their prediction model. A clear definition of the 35% threshold and its reason was explained. The VMT report describes the need for the assay in the current regulatory context (6). Furthermore, a clear rationale for the assay has been given (the rationale for the test method is that drugs and chemicals, environmental contaminants, food additives, and drugs can target the immune system, resulting in immune dysregulation). It is known that IL-2 exerts pleiotropic actions on CD4+ T cell differentiation via its modulation of cytokine receptor expression. It promotes Th1 differentiation by inducing IL-12Rb2 (and IL-12Rb1), promotes Th2 differentiation by inducing IL-4Ra, inhibits Th17 differentiation by inhibiting gp130 (and IL-6Ra), and drives Treg differentiation by inducing IL-2Ra. IL-2 also potently represses IL-7Ra, which decreases survival signals that normally promote cell survival and memory cell development (7). Therefore, it is reasonable that the test developer focused on the regulation of IL-2 transcription and attempted to construct an AOP of immunotoxicity with transcriptional dysregulation of IL-2 as a central key event. The VMT report mentioned that IL-2 Luc assay will be part of a broader tiered approach to eventually include IL-8 and IL-1β, that corresponds to the AOPs for immunotoxicity and as such constitutes an IATA.

The PRP agreed that the mechanistic basis of the method and how it related to the T-cell specific endpoint also was well described in the VMT report.

Within Laboratory Reproducibility

The PRP agreed that the results which emerged have demonstrated a sufficient degree of within laboratory reproducibility. For achieving such conclusion, the PRP focused on results obtained with the final protocol and prediction model.
A total of 5 coded chemicals (4 T-cell targeting and 1 non T-cell targeting) were evaluated by 3 experimental sets. Based on such assumptions, the success criterion of >80% within-laboratory reproducibility was achieved in each of the three participating laboratories (Lab. A: 80.0% (4/5), Lab. B: 100% (5/5), Lab. C: 80.0% (4/5)).

The PRP notes that extensive documentation of within-laboratory reproducibility data for the final and all the development phases of the IL-2 Luc assay has been displayed in the VMT report and its appendices. Taken into account, the data lends support to the view that the assay has a sufficient level of reproducibility within laboratories.

Interlaboratory Transferability

The PRP noted that the technical transfer of the IL-2 Luc assay involved training and successful assessment of 3 experiments of 5 test substances (not blinded) by each of the participating laboratories. That work was prior to their approval to participate in the subsequent validation work.

Between Laboratory Reproducibility

With regard to between-laboratory reproducibility, the PRP recognized that the test results gave 80% (20/25) and met the success criterion of >80% between-laboratory reproducibility. The number of test chemicals is combined of the Phase I (5) and Phase II (20) .

Again, the PRP notes that extensive and transparent documentation of between-laboratory reproducibility data for all phases of the IL-2 Luc assay has been displayed in the VMT report and its appendices.

The PRP concluded that the assay demonstrated successful between-laboratory reproducibility.

Predictive Capacity

To determine the predictivity of the IL-2 Luc assay, it is crucial to understand the immunotoxic characteristics of chemicals used in the study. The PRP agreed that classification chemicals into those that affect T cell function (T cell-targeting chemicals, TTC) and those that do not directly affect T cell function (non-T cell-targeting chemicals, NTTC) was needed. The PRP confirmed the rationale for classifying immunotoxic chemicals are clearly described in the validation report.

Demonstration of a test method's performance should be based on the testing of representative, preferably coded, reference chemicals. The PRP concluded that the validation study used an appropriate level of test chemical coding to ensure fully blinded evaluation. With respect to chemical selection, the PRP confirmed that the criteria for chemical selection were clearly outlined. On the other hand, it should be noted that there is a question as to whether or not the number of true negatives (8/25) in the set was sufficient.

The immunotoxic characteristics of each chemical used in the Phase I and Phase II are

shown in the VMT report and based on the criteria total 25 chemicals were classified into 16 positives, 8 negatives and 1 unclassified. According to the classification, accuracy is 75% (18/24), specificity is 75% (6/8) and sensitivity is 75% (12/16). The PRP concluded the predictive capacity of the test is not sufficient to detect all immunotoxic chemicals if used as a stand-alone test.

The PRP basically agreed with the test developer's opinion that there are at least 2 reasons that the predictivity did not meet the success criterion. First, the classification of immunotoxic chemicals is sometimes uncertain, because mechanistic information is often limited. Second, the IL-2 Luc assay does not cover every aspect of the effects of chemicals on T cell function.

Regarding the second point, the PRP noted again this assay should be used in combination with other assays, as in the context of IATA. With respect to the first point, the PRP noted the use of compounds with a better-defined immunotoxicity mechanism will help to ultimately improve the accuracy of the validation study and also recommended a Detailed Review Paper (DRP) on in vitro immunotoxicity testing.

Following phases I and II, the assay was then applied to over 60 chemicals that were previously evaluated by the test developer (8). The accuracy was calculated as 82.4%, which is in line with results obtained in phases I and II.

Applicability Domain

The PRP concluded that the applicability domain should be better defined. First, the method cannot detect immunotoxicity associated with inhibition of DNA synthesis and cell division. Second, the assay might not detect compounds that require metabolic activation to a toxic intermediate. Third, the use of PMA/Io as a stimulant bypasses signaling through the T cell receptor. Finally, the IL-2 Luc assay shares limitations common to many suspension cell-based techniques when testing highly hydrophobic substances.

Performance Standards

The PRP concluded that the list of performance standard (PS) substances placed in Appendix 16 to the VMT report was satisfactory. Mutual Acceptance of Data (MAD) will only be guaranteed for test methods validated according to the PS, if these test methods have been reviewed and included in this Test Guideline by the OECD. The PS is

supplemented by a list of proficiency chemicals, listed in Appendix 15, to be used as a routine check on performance of the assay.

Additional Comments

The PRP concluded that the validation study management and conduct met the criteria set out in OECD GD 34 (2005). The PRP noted that the study was conducted not under GLP certification but in the spirit of GLP.

The PRP appreciated the transparency with which all the IL-2 Luc assay material was presented. The PRP noted that during the review they were able to access the raw data files associated with the IL-2 Luc assay development/validation work.

The PRP also noted that AOP networks and a DRP in this field are essential for the construction of an IATA.

Conclusions and Recommendations

The PRP concluded that, even though the predictive capacity was not sufficient to allow use as a stand-alone test, the IL-2 Luc assay validation has demonstrated that the method should be acceptable as a part of IATA for the predictive screening of T-cell targeted immunotoxicity.

Acknowledgements

The PRP is also grateful to the members of the VMT for supplying the report on the validation study of the IL-2 Luc Assay. This research was supported by the WHLW grant.

References

1) Adler et al. (2010) Alternative (non-animal) methods for cosmetics testing: current status and future prospects. Arch Toxicol. 2011 85(5):367-485

2) Corsini E and Roggen EL., 2017. Overview of in vitro assessment of immunotoxicity. Curr Opin Toxicol. 5: 13-18.

3) Saito, R., Hirakawa, S., Ohara, H., et al., 2011. Nickel differentially regulates NFAT and NF-kappaB activation in T cell signaling. Toxicology and applied pharmacology, 254, 245-255.
4) Takahashi, T., Kimura, Y., Saito, R., et al., 2011. An in vitro test to screen skin sensitizers using a stable THP-1-derived IL-8 reporter cell line, THP-G8. Toxicological sciences: an official journal of the Society of Toxicology 124, 359-369.
5) Kimura, Y., Fujimura, C., Ito, Y., et al., 2014. Evaluation of the Multi-ImmunoTox Assay composed of 3 human cytokine reporter cells by examining immunological effects of drugs. Toxicol In Vitro 28, 759-768.6) (WHO)/, W.H.O., Meeting, I.P.o.C.S.I.S., 2012. Guidance for immunotoxicity risk assessment for chemicals. WHO press.
7) Liao, W., Lin, J.X., Wang, L., et al., 2011, Modulation of cytokine receptors by IL-2 broadly regulates differentiation into helper T cell lineages. Nat Immunol. 12: 551-559.
8) Kimura, Y., Fujimura, C., Ito, Y., et al., 2018. Profiling the immunotoxicity of chemicals based on in vitro evaluation by a combination of the Multi-ImmunoTox assay and the IL-8 Luc assay. Arch Toxicol 92, 2043-2054.