TEXTE
10/2017

Assessment of environmental persistence: regulatory requirements and practical possibilities – available test systems, identification of technical constraints and indication of possible solutions

Final Report

For our Environment

Umwelt Bundesamt

TEXTE 10/2017

Environmental Research of the
Federal Ministry for the
Environment, Nature Conservation,
Building and Nuclear Safety

Project No. 54429
Report No. (UBA-FB) 002326

Assessment of environmental persistence: regulatory requirements and practical possibilities – available test systems, identification of technical constraints and indication of possible solutions

by

Stefan Gartiser
Hydrotox GmbH, Freiburg

Klaus Schneider
Forschungs- und Beratungsinstitut Gefahrstoffe (FoBiG) GmbH,
Freiburg

Markus A. Schwarz
Forschungs- und Beratungsinstitut Gefahrstoffe (FoBiG) GmbH,
Freiburg

Thomas Junker
ECT Oekotoxikologie GmbH, Flörsheim

On behalf of the German Environment Agency

Imprint

Publisher:
Umweltbundesamt
Wörlitzer Platz 1
06844 Dessau-Roßlau
Tel: +49 340-2103-0
Fax: +49 340-2103-2285
info@umweltbundesamt.de
Internet: www.umweltbundesamt.de

/umweltbundesamt.de
/umweltbundesamt

Study performed by:
Forschungs- und Beratungsinstitut Gefahrstoffe GmbH (FoBiG)
Klarastraße 63
79106 Freiburg

Hydrotox GmbH
Bötzinger Straße 29
79111 Freiburg

ECT Oekotoxikologie GmbH
Böttgerstraße 2 – 14
65439 Flörsheim

Study completed in:
March 2016

Edited by:
Section IV 2.3 Chemicals
Claudia Staude, Ulrich Jöhncke

Publication as pdf:
http://www.umweltbundesamt.de/publikationen

ISSN 1862-4804

Dessau-Roßlau, January 2017

The Project underlying this report was supported with funding from the Federal Ministry for the Environment, Nature Conservation, Building and Nuclear safety under project number 54429. The responsibility for the content of this publication lies with the author(s).

Kurzbeschreibung

Um die Anwendbarkeit etablierter und neuer Bioabbaubarkeitstests für die Persistenzbewertung im Rahmen der PBT/vPvB-Beurteilung unter REACH zu prüfen, wurde eine Literaturstudie durchgeführt. Screeningtests auf leichte und inhärente biologische Abbaubarkeit, Vorschläge für *Enhanced Screening*-Tests sowie erst kürzlich entwickelte kompartimentspezifische Screening-Tests wurden hinsichtlich ihrer Stärken und Schwächen beurteilt. Darauf basierend werden Empfehlungen zur Verbesserung ihrer Leistungsfähigkeit und Eignung für die Persistenzbeurteilung gegeben. Obgleich nicht im Zentrum dieser Studie, werden auch für Simulationstests (OECD 307, 308, 309) Vorschläge zur Verwendung bei der Persistenzbewertung gemacht.

Die Empfehlungen konzentrieren sich auf die Definition, Standardisierung und Optimierung von Testbedingungen, Validitätskriterien und Auswertung im Hinblick auf die Persistenzbewertung. Bezüglich *Enhanced Screening*-Tests wird kritisch diskutiert, welche „Verstärkungen" eingeführt werden können, ohne den Screening-Charakter dieser Testgattung in Frage zu stellen. Darüber hinaus werden spezielle Problematiken im Rahmen der Persistenzbeurteilung adressiert, namentlich nicht-extrahierbare Rückstände (NER) in Tests mit Boden oder Sediment sowie Substanzgruppen, die in standardisierten Tests zu Problemen führen können (schwer wasserlösliche Substanzen; Substanzen hoher Flüchtigkeit; UVCB-Stoffe).

Abstract

A literature study was performed to review the applicability of established and new tests for biodegradability for assessing persistence in the frame of PBT/vPvB assessments under REACH. Screening tests for ready and inherent biodegradability, proposals for enhanced ready tests as well as newly designed compartment-specific screening tests were analysed for strengths and weaknesses and proposals are made how to improve their performance and suitability for assessments of persistence. Although not in the focus of this study, some recommendations are also given for simulation tests (OECD 307, 308, 309) in the context of evaluating persistence.

Recommendations focus on defining, standardising and optimising test conditions, on validity criteria and interpretation of test results. In the case of enhanced screening tests it is critically discussed which test modifications could be introduced without challenging the screening nature of the tests. Furthermore, specific issues such as non-extractable residues (NER) in tests with soil and sediment and substances difficult to test in standard tests (poorly water soluble, highly volatile or UVCB substances) are addressed.

Table of contents

List of figures

List of tables

Acronyms and Definitions

Adaption / adapted inoculum	Incubation of the inoculum with the test item before the start of the actual test – adapted inoculum is used to shorten the adaption phase or to reduce inoculum toxicity of the test item (e.g. by exposure to the test item at low concentrations) – see *pre-exposure*
ASTM	American Society for Testing and Materials
ATP	Adenosine triphosphate
BOD	Biological oxygen demand; BOD5 = BOD after 5 days
CEFIC	European Chemical Industry Council
COD	Chemical oxygen demand
DOC	Dissolved organic carbon
ECETOC	European Centre for Ecotoxicology and Toxicology of Chemicals
ECHA	European Chemicals Agency
EU	European Union
ISO	International Organization for Standardization
MLR	Mass loading rate, measured as the biological oxygen demand (BOD5) per total suspended solids per day (equivalent to the sludge loading rate, SLR)
MITI	Ministry of International Trade and Industry (Japan)
NADH	Nicotinamide Adenine Dinucleotide
NER	Non-extractable residues
OECD	Organisation for Economic Co-operation and Development
PAH	Polycyclic aromatic hydrocarbons
Pre-exposure	Any pre-treatment of the inoculum in presence of the test substance with the aim to obtain (pre-)adapted inoculum – see adaption
Pre-incubation	Incubation of inoculum at the test conditions, without the test item, before the start of the test: often performed to reduce background respiration of inoculum controls
REACH	Registration, Evaluation, Authorisation and Restriction of Chemicals (Regulation (EC) No 1907/2006)
RIVM	Dutch National Institute for Public Health and the Environment
SCAS	Semi-continuous activated sludge
SETAC	Society for Environmental Toxicology and Chemistry
SRT	Sludge retention time
STP	Sewage treatment plant
ThCO$_2$	Theoretical carbon dioxide evolution
ThOD	Theoretical oxygen demand

TNO	Toegepast Natuurwetenschappelijk Onderzoek (The Netherlands Organisation for Applied Scientific Research)
TOC	Total organic carbon
US EPA	United States Environmental Protection Agency
UVCB	Substances of Unknown or Variable Composition, Complex Reaction Products and Biological Materials

Zusammenfassung

Zur Persistenzbewertung im Rahmen einer PBT/vPvB-Bewertung unter REACH werden gegenwärtig entweder sogenannte Screeningtests auf leichte biologische Abbaubarkeit (*ready biodegradability*, OECD 301-Serie) oder aber Simulationstests für Boden, Sediment und Oberflächenwasser (OECD 307, 308 oder 309) angewandt. Simulationstests werden unter umweltrelevanten Bedingungen durchgeführt, sind aber teuer und die Ergebnisse hängen von vielen Variablen ab, was die Interpretation solcher Tests oftmals erschwert. Screeningtests auf leichte biologische Abbaubarkeit sind im Rahmen von REACH bereits ab der untersten Tonnageschwelle erforderlich. Zwar weisen diese Tests artifizielle Testbedingungen auf, die nicht mit denen in Umweltkompartimenten verglichen werden können, diese Bedingungen sind jedoch stringent und die anzuwendende Interpretation konservativ. Daher werden Ergebnisse aus Tests auf leichte biologische Abbaubarkeit zur Bewertung der Persistenz im Rahmen der PBT/vPvB-Prüfung als sogenannte *Screeninginformation* akzeptiert.

Aufgrund der Stringenz und Konservativität dieser Tests müssen allerdings viele als nicht leicht abbaubar (*not readily biodegradable*) – und damit auf Screening-Ebene als persistent gewertete – Substanzen weitergehend untersucht werden. Daher wird versucht, Screeningtests derart weiterzuentwickeln, dass sie verlässliche Aussagen zur Persistenz zulassen - ohne den mit Simulationstets verbundenen Aufwand und die häufig damit verbundenen Schwierigkeiten der Interpretation. Im Zentrum dieser Aktivitäten stehen die sogenannten *Enhanced Screeningtests*.

Idealerweise sollten (neue) Tests vom Screening-Typ folgende Eigenschaften aufweisen:

- Sie sollten deutlich schneller durchzuführen und billiger sein als Simulationstests.
- Sie sollten verlässlich und reproduzierbar sein.
- Sie sollten glaubwürdige Ergebnisse liefern, und zwar unter Berücksichtigung der vielfältigen Umweltbedingungen (keine Falschnegativen bezüglich Persistenz; eine niedrige Anzahl Falschpositiver bezüglich Persistenz).
- Sie sollten Vorhersagen für die verschiedenen Umweltkompartimente zulassen.

Diese Literaturstudie möchte einen Überblick zu laufenden Entwicklungen auf dem Gebiet jener Bioabbautests geben, die für eine Persistenzprüfung geeignet sein könnten. In Abschnitt 3 werden Tests auf leichte biologische Abbaubarkeit, inhärente Abbaubarkeit, Enhanced Screeningtests, kompartimentspezifische Screeningtests und Simulationstests auf ihre Vor- und Nachteile bezüglich der Persistenzbeurteilung analysiert. Als Basis für die Beurteilung von Relevanz und Anwendbarkeit dieser methodischen Ansätze werden die regulatorischen Anforderungen an die Persistenzbeurteilung in Abschnitt 4 zusammengefasst. Abschnitt 5 adressiert sowohl identifizierte Defizite als auch Möglichkeiten der Verbesserung für diese fünf Testkategorien. Schließlich werden in Abschnitt 6 Schlussfolgerungen bezüglich der Eignung vorhandener sowie vorgeschlagener Testsysteme für die Persistenzbeurteilung gezogen.

Screeningtests auf leichte biologische Abbaubarkeit: Defizite und Möglichkeiten ihrer Verbesserung

Die Entwicklung standardisierter Tests auf leichte biologische Abbaubarkeit begann in den 70er Jahren des vergangenen Jahrhunderts in verschiedenen Laboren, sodass ein jeder Test seine eigene Historie aufweist. Erst später wurden die unterschiedlichen Tests im Rahmen der OECD 301 (A-F) zusammengefasst und schließlich noch durch OECD 310 sowie ISO-Standards erweitert. Stoffe, die diesen Tests zufolge das Kriterium für leichte biologische Abbaubarkeit erfüllen, werden auch im Hinblick auf ihren Abbau in der Umwelt als (zumindest unter den meisten Bedingungen) schnell abbaubar angesehen und daher als in der Umwelt nicht persistent betrachtet. Dafür müssen sie die in diesen Tests gültigen Prüfkriterien erfüllen (sog. *pass levels*). Diese Prüfkriterien sind – je nach Testart - 60% Abbau gemessen am theoretischen Sauerstoffbedarf (ThOD) bzw. der theoretischen CO_2-Bildung

($ThCO_2$) sowie 70% bezüglich des gelösten organischen Kohlenstoffs (DOC). Diese Ergebnisse sind innerhalb einer Zeitdauer von 28 Tagen zu erzielen, mit der Besonderheit, dass im Hinblick auf die Persistenzbewertung das sonst einzuhaltende 10-Tage-Fenster (gemessen ab 10% Abbau bis zum Erreichen des Prüfkriteriums) nicht relevant ist. Diese Testmethoden zielen *nicht* darauf ab, Geschwindigkeitskonstanten für den Abbau zu ermitteln, sondern vielmehr, unter definierten Bedingungen und gemessen an der Ausgangskonzentration einen Mindestabbau innerhalb einer begrenzten Zeitspanne (28 d) zu erreichen und daran die leichte biologische Abbaubarkeit zu messen.

Die Analyse der bestehenden Testsysteme hat folgende Defizite hinsichtlich Genauigkeit, Reproduzierbarkeit und Vergleichbarkeit ergeben:

- Nach OECD 301 sind unterschiedliche Inokulumtypen (auch Gemische daraus) erlaubt, während die maximale Inokulumdichte durch die jeweiligen (testspezifischen) Validitätskriterien für die Inokulum-Hintergrundwerte („inoculum blank“) bestimmt werden (z.B. OECD 301 B max. 40 mg /L CO_2 innerhalb 28 Tagen). Ohne den prinzipiellen Charakter der leichten biologischen Abbaubarkeit in Frage zu stellen, könnten Testlabore die Empfindlichkeit ihrer Tests steigern, indem sie stets Inokulumdichten wählten, die nahe an die jeweiligen Validitätskriterien reichen und / oder das Flaschenvolumen erhöhten.
- Während die praktisch maximal mögliche Inokulumdichten in den Tests auf leichte biologische Abbaubarkeit (OECD 301) durch die erwarteten Inokulumhintergrundwerte limitiert wird (entsprechend der Validitätskriterien), sind diese Validitätskriterien zwischen den einzelnen Tests nicht konsistent. Eine systematische Definition dieser Obergrenzen fehlt bislang. Zudem solle das verwendete Inokulum besser beschrieben werden, beispielsweise durch die sogenannte „mass loading rate“, gemessen als der biologische Sauerstoffbedarf je suspendierter Gesamtschwebstoffe pro Tag, für den jeweils verwendeten Belebtschlamm. Auch andere mikrobiologische Parameter sind denkbar, um durch eine so erzielbare bessere Charakterisierung des Inokulums zu einer besseren Reproduzierbarkeit von Testergebnissen zu kommen.
- Während höhere Inokulumkonzentrationen durch eine oftmals praktizierte Vorbehandlung / Vorinkubation des Inokulums möglich werden, haben solche Ansätze Nachteile, da in der Regel die mikrobiologische Potenz durch solcherlei Maßnahmen reduziert wird. Derzeit gibt es keine vielversprechenden Ansätze einer weitergehenden Standardisierung des Inokulums, beispielsweise im Sinne von definierten bakteriellen Stämmen.
- Da viele jener Studien, die den Einfluss der Quelle, Qualität, Vorbehandlung, Konzentration und absolute Menge des Inokulums untersucht haben, nicht unter OECD 301-konformen Testbedingungen durchgeführt wurden, bleibe die tatsächliche Konsequenz dieser Parameter auf die Testergebnisse von „Ready“-Tests oder *Enhanced Screeningtests* unklar.
- Die geforderte Replikatezahl in den OECD 301-Tests ist gering– dies schließt die Replikate für die Inokulumhintergrundatmung (“inoculum blanks”) ein (beispielsweise nur n=1 für die Inokulumhintergrundatmung im MITI-I-Test). Entsprechend sollten Validitätskriterien für die Variabilität der Inokulumhintergrundatmung in Parallelinkubationen entwickelt werden. Gegenwärtig sieht OECD 301 lediglich eine obere Grenze von 20% für die Variabilität des Abbaus in parallelen Testgefäßen vor, jedoch ohne diese in den Inokulumkontrollen zu begrenzen.
- Im Falle von DOC-basierten Tests wie dem OECD 301 A sollte der Einfluss der Adsorption an den Belebtschlamm auf das Testergebnis gründlich untersucht werden: Momentan existiert kein Leitwert, bis zu welchem Maße eine adsorptionsbasierte Eliminierung noch tolerable wäre. Die Einführung eines entsprechenden Validitätskriteriums verbunden mit abiotischen Kontrollen wird als notwendig erachtet.
- Ausschließlich auf der Wasserphase basierende Testsysteme berücksichtigen nur sehr unzureichend Prozesse wie Adsorption und Desorption, die aber sowohl auf Bioverfügbarkeit wie Abbau Einfluss nehmen können.

- Für schlecht wasserlösliche Substanzen sollten spezielle Referenzsubstanzen schlechter Wasserlöslichkeit *und* bekannter Bioabbaubarkeit entwickelt werden. Dies würde Verlässlichkeit und Vergleichbarkeit von Testergebnissen für diese schwierige Substanzgruppe erhöhen. Darüber hinaus müssen die Anwendbarkeitskriterien der einzelnen Tests für flüchtige Substanzen klar definierte werden.

Reproduzierbarkeit, Verlässlichkeit und Vergleichbarkeit (zwischen unterschiedlichen Tests, aber auch zwischen Testergebnissen zum gleichen Test aus unterschiedlichen Laboratorien) könnten durch die genannten Verbesserungen erhöht werden. Und schließlich könnte die Anwendbarkeit von Tests zur leichten biologischen Abbaubarkeit für die Persistenzprüfung verbessert werden, indem durch diese Vorschläge eine höhere Sensitivität erreicht wird, also eine Reduzierung der Quote von fälschlicherweise als nicht abbaubar eingestuften Stoffen.

Enhanced Screeningtests für die Bioabbaubarkeit: Mögliche Modifikationen („*enhancements*“) und ihre Anwendbarkeit für die Persistenzbewertung

Enhanced Screening-Tests werden für die Persistenzbewertung entwickelt, nicht für die Bewertung der leichten biologischen Abbaubarkeit. Ihr Ziel ist es, die Durchführung aufwendiger Simulationstests zu vermeiden. Folgende Modifikationen auf Grundlage der Tests zur leichten biologischen Abbaubarkeit werden gegenwärtig diskutiert: die Verlängerung des Tests über 28 Tage hinaus; die Verwendung größerer Testgefäße bei gleicher Inokulumkonzentration; die Erhöhung der Inokulumkonzentration; eine Präexposition gegenüber der geringkonzentrierten Testsubstanz (Adaptation). Wie schon bei den Tests auf leichte biologische Abbaubarkeit ausgeführt, wird auch hier der Inokulumquelle sowie der Inokulumqualität ein hoher Stellenwert eingeräumt. Die Vergrößerung der Testgefäße wie auch die Steigerung der Inokulumkonzentration erhöhen die Wahrscheinlichkeit, abbaukompetente Mikroorganismen mit dabei zu haben. Während (anders als bei den Tests auf leichte biologische Abbaubarkeit) nach REACH Leitfadendokument R.7B die Einhaltung eines 10-Tage-Fensters hier nicht relevant ist, sollten die entsprechenden Prüfkriterien ($\geq$ 60% ThOD /$ThCO_2$ bzw. $\geq$ 70% DOC) auch für Enhanced Screening-Tests zur Persistenzbeurteilung herangezogen werden. Auch für diese Tests besteht nicht die Absicht, kinetische Abbauraten zu bestimmen, da die Testbedingungen nicht den Bedingungen in natürlichen Umweltkompartimenten entsprechen.

Nach REACH Annex XIII umfassen die für die Persistenzbewertung heranzuziehenden Informationen auch Screeningtests, explizit auch *Enhanced Screening*-Tests sowie andere Informationen, vorausgesetzt „...Eignung und Zuverlässigkeit [können] angemessen nachgewiesen werden...“. REACH Leitfadendokument R.7B schränkt aber ein, dass nur *Enhanced Screening*-Tests mit Inokulum, welches ***nicht*** aus dem Kläranlagenumfeld stammt, für die Persistenzbewertung (um Nicht-Persistenz zu zeigen) herangezogen werden sollten. Da die allermeisten Tests auf leichte biologische Abbaubarkeit mit kläranlagenstämmigen Inokula durchgeführt werden, und eines der verbreitetsten „Enhancements“ die Verlängerung über 28 Tage hinaus darstellt, würde diese Einschränkung nach R.7B die Anwendbarkeit des Test-Typs *Enhanced Screening* wesentlich einschränken. Darüber hinaus fehlt die wissenschaftliche Rechtfertigung dafür: Kläranlagenstämmige Inokula sind für Tests auf leichte biologische Abbaubarkeit zugelassen und die Tests werden für die Persistenzbewertung unter REACH akzeptiert; kläranlagenstämmige Inokula werden über den Kläranlagenausfluss in Umweltgewässer eingetragen (einschließlich Schwebstoffe); kläranlagenstämmige Inokula führen generell zu höherer Reproduzierbarkeit verglichen mit umweltstämmigen Inokula, beispielsweise Oberflächenwasser oder filtrierte Bodeneluate, die starken lokalen und saisonalen Schwankungen unterliegen. Solange keine Präadaption an die Testsubstanz vorliegt, wird daher die Verwendung kläranlagenstämmigen Inokulums in *Enhanced Screening*-Tests zur Persistenzbeurteilung als vertretbar angesehen. Betont werden muss hier, dass weder Tests auf leichte biologische Abbaubarkeit noch *Enhanced Screening*-

Tests für sich beanspruchen, reale Umweltbedingungen anzunähern oder gar zu repräsentieren; vielmehr wurden sie entwickelt, um generell anwendbare Aussagen zur biologischen Abbaubarkeit einer Substanz zu erzielen.

Ein weiteres Hemmnis für die Anwendung von *Enhanced Screening*-Tests in der Persistenzbewertung besteht in ungenügenden Leitlinien bezüglich einerseits der Kombinierbarkeit von „Verstärkungen" und andererseits der Limitierung ihres Ausmaßes. Vor dem Hintergrund von verlässlichen und geeigneten Informationen, die für die Screeningbewertung der Persistenz nach Annex XIII REACH benötigt werden, bedarf es daher einer kritischen Hinterfragung zur Akzeptierbarkeit möglicher Modifikationen. Unsere Analyse erbrachte die folgenden Ergebnisse:

- Der Begriff *Enhanced Screening*-Tests ist innerhalb der REACH Leitfadendokumente nicht klar definiert: Zum Teil wird zwar der Begriff „enhanced ready biodegradation test " benutzt (Dokument R.11), eine klare Aussage dahingehend, dass es sich ausschließlich um eine Abwandlung von Tests auf leichte biologische Abbaubarkeit handelt und es damit "enhanced inherent screening tests" nicht geben kann, fehlt jedoch und sollte erfolgen. Entsprechend sollte die maximal akzeptierbare Inokulumkonzentration in *Enhanced Screening*-Tests unter derjenigen aus Tests auf inhärente biologische Abbaubarkeit liegen, um den Charakter des zugrundeliegenden Tests beizubehalten.
- Inokula von kontaminierten Standorten oder solche, die mit der Testsubstanz vorexponiert wurden, werden als ungeeignet für eine prospektive Persistenzbewertung angesehen und sollten daher in solchen Tests nicht verwendet werden.
- Eine Verlängerung der Testdauer auf bis zu 60 Tage wird als nicht kritisch angesehen: Eine derartige Testverlängerung ist bereits nach OECD und den REACH Leitfadendokumenten vorgesehen für den Fall, dass die Abbaukurve am Ende des Tests noch kein Plateau erreicht haben sollte. Allerdings können positive Abbauresultate aus solchen Tests nur zur Zurückweisung eines Persistenzverdachts verwendet, nicht aber zur Demonstration leichter biologischer Abbaubarkeit herangezogen werden.
- Auch die Vergrößerung des Testvolumens ist akzeptabel, da bereits möglich unter OECD 301, die lediglich Vorschläge, aber keine Begrenzung für diese Größe vorsieht. Die praktische Durchführbarkeit setzt dieser Modifikation eine natürliche Grenze.
- Derzeit gibt es keine Überlegungen dahingehend, Co-metabolismus – für bestimmte Verbindungen ein wesentlicher Abbaumechanismus – in die Konzeption von *Enhanced Screening*-Tests einzuschließen: Dies ist bedingt durch die grundlegende Testkonzeption, die auf der Messung von unspezifischen Summenparametern (DOC-Eliminierung, Sauerstoffzehrung oder CO_2-Entwicklung) beruht, wodurch die Einführung weiterer Kohlenstoffquellen einen wesentlichen Unsicherheitsfaktor bedeuten würde.
- Bei der Durchführung von *Enhanced Screening*-Tests sollte erwogen werden, zusätzliche positive und negative Referenzsubstanzen mit einzuschließen: Diese Substanzen sollten entsprechend der Komplexität des Testsystems gewählt und beispielsweise abbaubar sein unter „enhanced"-Bedingungen, aber nicht leicht biologisch abbaubar sein (Positivkontrolle) bzw. auch unter „enhanced"-Bedingungen nicht wesentlich abgebaut werden (Negativkontrolle).
- Prüfkriterien für *Enhanced Screening*-Tests sind bislang nicht klar definiert: Zunächst sollten daher die Prüfkriterien nach OECD 301 beibehalten werden, wie sie derzeit für die Persistenzbeurteilung heranzuziehen sind (siehe oben).
- Derzeit existieren keine für die *Enhanced Screening*-Tests spezifischen Validitätskriterien, während die in OECD 301 spezifizierten Kriterien aufgrund der eingebrachten „Enhancements" zumindest teilweise nicht mehr anwendbar sind. Grundsätzlich sollten die Unterschiede zwischen den Messungen der Hintergrundreihe (Inokulum ohne Testsubstanz) und der Testreihe eine möglichst genaue Abschätzung des Abbaugrades ermöglichen. Dies kann wesentlich verbessert werden, wenn zusätzliche Replikate sowohl für die Test- als auch die Hintergrundreihen eingeführt

werden, wodurch eine genauere Abschätzung der Variabilität und damit der Genauigkeit erfolgen kann.

Es bleibt zu klären, welche "Verstärkungen" (*enhancements*) als solche und in welcher Kombination zu Resultaten führen, die Schlüsse zum Abbauverhalten unter Umweltbedingungen (und damit zur Persistenz) erlauben und somit aus regulatorischer Perspektive zulässig sind. Prinzipiell können alle Optimierungen einbezogen werden, die im Sinne der leichten biologischen Abbaubarkeit (*ready biodegradability*) unproblematisch sind (siehe oben), darunter die Vergrößerung des Testvolumens, die Verwendung gemischter Inokula (innerhalb der Validitätskriterien), sowie die Einführung zusätzlicher Replikate und/oder zusätzlicher Positiv- und Negativkontrollen. Was allerdings kritisch hinterfragt werden muss, ist die Kombination von eigentlichen Verstärkungen, also Modifikationen über diese genannten, im Rahmen der Tests zur leichten biologischen Abbaubarkeit liegenden, Optimierungen hinaus (z.B. Erhöhung der Inokulumdichte). Zwar intendieren Screeningtests nicht, Umweltbedingungen zu repräsentieren. Allerdings erfordert die Nutzung von Ergebnissen aus diesen Tests für die Persistenzbewertung, dass das Extrapolationspotential von diesen artifiziellen Bedingungen zu Umweltbedingungen für in den Testbedingungen vorgenommene Veränderungen („Enhancements") vorher entsprechend validiert wurde. Wir schlagen daher Folgendes vor:

1. Für die jeweils möglichen "Verstärkungen" sind klare Grenzen zu setzen (siehe oben): beispielsweise die Begrenzung der Inokulumkonzentration auf Werte unterhalb jener für Inhärenttests; Ausschluss vorexponierten Inokulums; Begrenzung einer möglichen Testverlängerung auf 60 Tage; Wahl des Testvolumen nach Belieben.
2. Die Auswirkung von Kombinationen von "Verstärkungen" sollten evaluiert werden, und es muss ein Ausschluss solcher Kombinationen erfolgen, die den Screening-Charakter (hohes Extrapolationspotential zu Umweltbedingungen) der *Enhanced Screening*-Tests beeinträchtigen.
3. Prinzipiell hat, basierend auf Substanzen bekannter biologischer Abbaubarkeit, eine kritische Diskussion der Leistungsfähigkeit solcher *Enhanced Screening*-Tests zu erfolgen. Basierend auf Ergebnissen mit solchen Substanzen sollten geeignete Prüfkriterien („pass levels") und Validitätskriterien etabliert werden.

Screening-Tests auf inhärente Abbaubarkeit: Defizite, mögliche Verbesserungen und Anwendbarkeit für die Persistenzbewertung

Sogenannte Inhärenttests werden unter für die biologische Abbaubarkeit sehr günstigen Umständen durchgeführt und beantworten daher die ebenfalls relevante Frage, ob eine Substanz überhaupt und prinzipiell ein Potenzial für die vollständige Mineralisierung durch biologische Abbauvorgänge besitzt, unabhängig davon, wie relevant dieses Ergebnis im Einzelfall für ein Umweltkompartiment sein mag. Ergebnisse aus solchen Tests können in zweierlei Art genutzt werden: Einmal kann für eine Substanz der Persistenzverdacht zurückgewiesen werden, falls ein Abbau $\geq$ 70% erreicht und zusätzlich spezifische Kriterien erfüllt wurden (z.B. für OECD 302 B: log-Phase nicht länger als 3 Tage, Prüfkriterium innerhalb von 7 Tagen erreicht), was zusammen als starker Indikator für Mineralisierung auch unter Umweltbedingungen angesehen wird. Zum anderen können negative Resultate aus diesen Tests (< 20% DOC-Eliminierung) als zuverlässiger Indikator für die Persistenz der Verbindung unter Umweltbedingungen gewertet werden.

Für Inhärenttests wurden folgende Probleme und Möglichkeiten zur Verbesserung festgestellt:

- Derzeit existiert kein standardisierter inhärenter Abbautest auf Basis von CO_2-Entwicklung, obgleich in der Literatur einige solcher auf den OECD Testrichtlinien 301B und 310 basierenden

Tests beschrieben sind. Dabei hat insbesondere die Kombination der Endpunkte DOC- Eliminierung und CO_2-Bildung das Potential, vergleichsweise präzise zwischen adsorptionsbedingter Eliminierung und Mineralisierung zu differenzieren.

- Der MITI II-Test (OECD 302C) verlangt die Verwendung einer sehr spezifischen Mischung verschiedener Inokula, verbunden mit einer Präinkubation – ebenso akzeptabel wäre aber Inokulum aus anderen Quellen, beispielsweise Belebtschlamm oder einer andersgearteten Mischung von Inokula unterschiedlicher Umweltkompartimente.
- Bei der Anwendung DOC-basierter Tests (wie dem Zahn-Wellens Test nach OECD 302 B) muss die Möglichkeit der Elimination durch Adsorption sorgfältig untersucht werden – der 3-Stundenwert im Zahn-Wellens-Test sollte daher stets berichtet werden, auch wenn kein spezieller Verdacht auf Adsorption besteht.

Allerdings wird die Anwendbarkeit von Tests auf inhärente Abbaubarkeit für die Persistenzbeurteilung aufgrund der zu erfüllenden spezifischen Kriterien (vgl. Abschnitt 4.2.2) immer beschränkt bleiben, wenn es darum geht, den Persistenzverdacht zurückzuweisen. Dies ist in Anbetracht der hier hohen Inokulumkonzentration, die einzig für Kläranlagen repräsentativ sein kann, auch gerechtfertigt. Gleichzeitig impliziert dies, dass für *Enhanced Screening*-Tests die Steigerung der Inokulumkonzentration eine obere Grenze haben muss, die deutlich unterhalb derer liegt, die in Inhärenttests Verwendung findet (beispielsweise < 200 mg/L Belebtschlammtrockengewicht).

Trotz der genannten möglichen Verbesserungen ist offensichtlich, dass Inhärenttests kein essentieller Bestandteil einer generischen Strategie zur Persistenzbewertung sein werden, sie können jedoch in speziellen Fällen relevante Informationen liefern.

Medienspezifische Screeningtests: Testkonzeption, Defizite und Anwendbarkeit für die Persistenzbewertung

Medienspezifische Screeningtests integrieren Prozesse, die Bioverfügbarkeit und Abbau beeinflussen können (z.B. Sorption), und generieren damit kompartimentspezifisch Informationen zur Abbaubarkeit. Die Möglichkeit vorausgesetzt, aus solchen Tests relevante Halbwertszeiten (apparent erster Ordnung) ableiten zu können, wären solcherlei Tests bezüglich der kompartimentspezifischen Persistenzbewertung den entsprechenden Simulationstests am nächsten. Allerdings verhindert gerade die hohe Testmaterialkonzentration, die aus nachweistechnischen Gründen (keine Radioisotopmarkierung) in solchen Tests eingesetzt werden muss, die Ableitung von unter Umweltbedingungen relevanten kinetischen Daten. .

Einige publizierte, bislang aber nicht standardisierte medienspezifische Screeningtests wurden entwickelt. Man kann annehmen, dass diese Tests realitätsnähere Ergebnisse liefern werden als Screeningtests, die belebtschlammstämmige Inokula ohne weiter Zugabe von Feststoffen verwenden. Die Tests zeigen folgende Charakteristika:

- Adsorptionsprozesse werden, zumindest in einem gewissen Ausmaß und abhängig vom spezifischen Testaufbau, abgebildet.
- Entsprechend ist zu erwarten, dass Abbauraten (verglichen mit konventionellen Screeningtests) niedriger sind bei Substanzen mit relevantem Adsorptionspotenzial.
- Abbauraten werden stärker kompartimentspezfisch sein, soweit kompartimentspezifisches Inokulum Verwendung findet (was nicht für jedes Testsystem der Fall sein muss).

Um die Akzeptanz dieser Tests für die Persistenzbewertung zu erreichen, könnten folgende Maßnahmen notwendig werden:

- Die zulässige Testdauer (Zeit für den Abbau) muss möglicherweise nach oben angepasst werden (über 28 Tage hinaus), um ein Erreichen der einschlägigen Prüfkriterien für O_2-Verbrauch bzw. CO_2-Bildung (auch für als leicht biologisch abbaubar identifizierte Verbindungen) erreichen zu können.
- Andererseits könnte es notwendig sein, die gültigen Prüfkriterien für O_2-Verbrauch bzw. CO_2-Bildung selbst anzupassen (unter die gegenwärtig gültige Marke von 60% in "Ready"-Tests). Dies deuten erste Ergebnisse mit als leicht biologisch abbaubar (OECD 301) charakterisierten Stoffen (z.B. Anilin) an.

Bis dato fehlt für diese Tests eine fundierte Bewertung, und auch die einschlägigen Leitfäden der ECHA zu REACH geben hinsichtlich dieser vergleichsweise neuen Tests keine Hinweise. Prinzipiell können kompartimentspezifische Screeningtests zum biologischen Abbau (z.B. für Sediment und Boden) analog den Tests auf leichte biologische Abbaubarkeit verwendet werden, um eine schnelle Abbaubarkeit im entsprechenden Umweltmedium abzuleiten. Allerdings gibt es bislang wenig Erfahrung mit diesen Tests, und entsprechend besitzen sie noch keine generelle regulatorische Akzeptanz. Zur Anwendung im regulatorischen Kontext sind folgende Kriterien zu erfüllen:

1. Es bedarf der Entwicklung und Abstimmung eines konsolidierten Testaufbaus, der reproduzierbare und relevante Abbauergebnisse liefert und damit geeignet ist, in der Umwelt (potentiell) persistente Verbindungen zu identifizieren. Dies könnte durch eine Laborübergreifende Prüfung einer Testmethode erfolgen (sogenannter Ringtest).
2. Unter Berücksichtigung des Screening-Charakters dieser Tests werden weiterhin klare Kriterien für die Ergebnisinterpretation benötigt, die einen Stoff schließlich als in der Umwelt (potentiell) persistent oder eben nicht persistent einordnen. Solche sogenannten Vorhersagemodelle könnten auf Prüfkriterien basieren, die ggf. auch spezifische Abbauraten einschließen, müssen aber letztlich in einer Ja/Nein-Antwort bezüglich der Persistenz münden. Ein direkter Vergleich mit den Annex XIII-Kriterien aus REACH wird in der Regel abzulehnen sein, da die Testbedingungen nicht oder nur unzureichend die Umweltbedingungen spiegeln.
3. Sind schließlich Testaufbau und Auswertekriterien abgestimmt, sollte ein betreffender Test zu seiner Validierung auf eine definierte Auswahl von Verbindungen angewendet werden, deren Abbauverhalten bekannt und möglichst diversifiziert ist, um solche neuen bzw. veränderten Methoden zu validieren.
4. Die Validierung unter Anwendung dieses Test-Sets an Verbindungen sollte einerseits gegenüber Testergebnissen aus OECD 301-Tests erfolgen, andererseits – soweit verfügbar – gegen Ergebnisse aus gut etablierten Simulationsstudien (Literatur und Datenbanken) für das betreffende Kompartiment des medienspezifischen Screeningtests.

Zusammenfassend ist festzustellen, dass die in jüngerer Zeit entwickelten medienspezifischen Tests (für Boden und Sediment) ihrer Natur nach Screeningtests sind. Die Tests stellen potentiell eine vielversprechende Option zur Ermittlung kompartimentspezifischer Daten für jene Substanzen dar, für die aufgrund ihrer Stoffeigenschaften eine überwiegende Verteilung in das betreffende Kompartiment erwartet wird, bzw. wo im betreffenden Kompartiment ein vom pelagischen (oder sonstigen) Vergleichskompartiment abweichender Abbau (qualitativ oder quantitativ) zu erwarten ist. Die genaue Rolle dieser Tests innerhalb einer Teststrategie zur Persistenzbeurteilung ist noch zu bestimmen.

Simulationstests zum Abbau in Umweltmedien: Defizite und mögliche Verbesserungen

In den vergangenen Jahren wurden eine Reihe von Defiziten und Unzulänglichkeiten existierender Simulationstests festgestellt und diskutiert. Beim gut etablierten OECD 307 zur aeroben und anaero-

ben Transformation im Boden (Bodensimulationstest) betrifft dies vorrangig die Frage, wie nicht extrahierbare Rückstände (NER, für *non-extractable residues*) und gebundene Rückstände (BR, für *bound residues*) zu definieren, zu bestimmen und zu interpretieren sind.

Kritische Diskussionspunkte in Bezug auf den Wasser-Sediment-Simulationstest (OECD 308) sind:

- Da der kontinuierliche Austausch zwischen Wasser- und Sedimentphase im Testsystem weder standardisierbar noch quantifizierbar ist, ist es unmöglich, belastbare kompartimentspezifische Halbwertszeiten für den Abbau abzuleiten.
- Die Redoxbedingungen in der Sedimentphase sind nicht definiert, da im Test eine aerobe Wassersäule über einer dünnen aeroben Sedimentschicht simuliert wird, unterhalb derer sich der anaerobe Sedimentbereich befindet. Daher sind die konkreten Testbedingungen ganz wesentlich von Faktoren beeinflusst, die für die Sauerstoffverteilung im Sediment wesentlich sind, wie beispielsweise der Belüftungsrate, den Strömungsverhältnissen in der Wasserphase, sowie von Sedimentdicke und –textur.
- Das empfohlene Verhältnis von 3 bis 4 Teilen Wasser auf einen Teil Sediment zusammen mit der empfohlenen Sedimentstärke von 2,5 ± 0,5 cm stellt in Bezug auf die für Industriechemikalien charakteristischen Expositionsszenarien keine relevanten Umweltbedingungen dar: Aufgrund des vergleichsweise geringen Wasseranteils liegt die Massenverteilung eines Stoffs im Gleichgewicht oft auf der Seite der Sedimentphase.
- Der resultierende Bereich möglicher Systemgeometrien (z.B. die Höhe der Wasser und Sedimentschicht, die Größe der Phasengrenzfläche) kann sowohl Einfluss nehmen auf die Verteilungsprozesse (bestimmt durch den Gleichgewichtsverteilungskoeffizienten sowie der Diffusionsrate von der Phasengrenzfläche Wasser/Sediment in die Sedimentphase) wie auch die Bioabbauprozesse und damit auf die im experimentellen System beobachtete Persistenz. Damit sind die Ergebnisse in gewissem Maße Artefakte des jeweiligen konkreten Testsystems und die Übertragbarkeit auf die Umwelt ist problematisch.
- Auch der statische Testaufbau entspricht nicht üblichen Umweltbedingungen, da Einflussgrößen wie Strömungsgeschwindigkeit und Sedimentdynamik nicht abgebildet werden.
- Die aus diesem Test abgeleitete Halbwertszeit für das Gesamtsystem Wasser/Sediment) kann nicht direkt mit den Kriterien des Annex XIII verglichen werden (*separat* für Wasser und Sediment).

Bezüglich des dritten Simulationstest-Typs, dem Test zum Abbau in Oberflächengewässern (OECD 309) wurden kürzlich folgende Möglichkeiten der Verbesserung von Testleistung und Testauswertung vorgeschlagen:

- Die große Anzahl möglicher experimenteller Variationen bedingt eine große Variationsbreite der Testergebnisse. Es bedarf daher der Spezifizierung des Testaufbaus (z.B. bzgl. der gewählten Temperatur (z.B. 12 oder 20°C)). Das Testsystem sollte geschüttelt und nicht gerührt werden, um eine Zerkleinerung des Sediments (dessen Beimischung eine der möglichen Optionen darstellt) und eine oft damit einhergehende verstärket Adsorption zu vermeiden.
- Die vorgesehene Option, diffuse zu beleuchten, bedarf der Ausarbeitung: Der Licht-Dunkelzyklus, die Lichtintensität sowie die Wellenlängenverteilung müssen klar definiert werden.
- Neue Modellierungsansätze könnten die Vorhersage eines kompartimentungebundenen Bioabbauparameters, namentlich der bioverfügbarkeitskorrigierten, Biomasse-normalisierten Bioabbaurate zweiter Ordnung (k'_{bio}) erlauben. Diese Größe erlaubt es, die Persistenz an der Phasengrenze Wasser-Sediment zu beurteilen und wäre damit auch relevant für den Sediment-Simulationstest (OECD 308).

Schließlich lässt sich sagen, dass Simulationstests aufgrund ihres hohen Zeit- und Ressourcenbedarfs gegenwärtig nur als letzte Stufe beim Test der biologischen Abbaubarkeit eingesetzt werden. Obwohl sie nicht im Zentrum dieser Literaturstudie standen, konnten einige Möglichkeiten zu ihrer Verbesserung identifiziert werden. Der aktuelle Stand ist im Folgenden kurz zusammengefasst:

1. Gut etabliert ist der Simulationstest zum Bioabbau im Boden nach OECD 307. Wie für jeden Bodentest stellt die Identifizierung und Bewertung von nicht extrahierbaren Rückständen eine Herausforderung dar an der weiter gearbeitet werden muss.
2. Sowohl OECD 308 (Wasser-Sedimenttest) als auch OECD 309 (Oberflächenwassertest) würden von einer Spezifizierung sowohl des Testaufbaus als auch der Kriterien zur Testdurchführung profitieren.

Schwer wasserlösliche Stoffe

Prinzipiell können auch schwer wasserlösliche[1] Stoffe bereits gegenwärtig in Screening-Tests auf ihre Abbaubarkeit hin untersucht werden, insbesondere, wenn dafür vorgesehene spezielle Modifikationen vorgenommen werden. Allerdings besteht im Vergleich zu löslichen Verbindungen das Risiko die Bioabbaubarkeit zu unterschätzen, da die in Screening-Tests effektiv verfügbare Konzentration niedrig sein kann (begrenzt durch die thermodynamisch bedingte Löslichkeit, kinetischer Hemmung des Lösungsvorgangs oder der zugänglichen Oberfläche). Als Konsequenz daraus könnte in extremen Fällen ein relevantes Biomassewachstum ausbleiben. Da aber gerade das Biomassewachstum die charakteristische sigmoidale Abbaukurve bedingt, die typisch für diese Tests ist, kann die Bioabbaubarkeit stark unterschätzt werden. Demgegenüber kann die viel geringere Nominalkonzentration, die üblicherweise in Simulationstests eingesetzt wird, sogar für schwerlösliche Verbindungen nahe oder unterhalb ihrer Löslichkeitsgrenze liegen, so dass ein effektiver Bioabbau beobachtet werden kann (vorausgesetzt, die Verbindung ist nicht prinzipiell abbauresistent).

Um Leistungsfähigkeit und Interpretation von Tests zur leichten biologischen Abbaubarkeit bzw. *Enhanced Screening*-Tests für diese Substanzgruppe zu verbessern, sollten (zusätzliche) Referenzverbindungen geringer Wasserlöslichkeit, aber hinreichender Bioabbaubarkeit (z.B. mikrokristalline Zellulose) validiert und im Test mitgeführt werden.

UVCB und aus mehreren Bestandteilen bestehende Substanzen (*multi-constituent substances*, MCS)

Ein systematisches und praktisch anwendbares Konzept für die Persistenzbewertung von UVCB (*Substances of Unknown or Variable Composition, Complex Reaction Products and Biological Materials*) fehlt bislang:

Während im Rahmen des PBT/vPvB-Assessments unter REACH eine Prüfung der Persistenz von Einzelbestandteilen einer Substanz bis zu einer Konzentration von 0,1 Gewichtsprozent zu erfolgen hat, müssen für UVCB nur Einzelbestandteile ab 10% spezifiziert werden. Andere Bestandteile sind oft ungenügend definiert, was sowohl eine radioaktive Markierung als auch die Anwendung von QSAR-Modellen praktisch unmöglich macht. Somit stellen Screening-Tests zur Bioabbaubarkeit den einzigen praktikablen Weg dar, einen Persistenzverdacht zurückzuweisen. Allerdings wird deren Anwendbarkeit durch das einschlägige Leitfadendokument zu REACH auf UVCB begrenzt, die aus homologen Strukturen bestehen. Für andersartige UVCBs sollte die Bioabbaubarkeit auf Basis eines fallspezifi-

[1] Gegenwärtig gibt es keine einheitliche Definition schwerer Wasserlöslichkeit: REACH Leitfaden R.7b verweist auf OECD (< 100 mg/L) und EU-TGD (2003) (< 1 mg/L) und merkt an, dass eher unterhalb von 1 mg/L damit zu rechnen ist, dass in Tests auf Umweltverhalten und Ökotoxizität Probleme auftreten werden.

schen Vorgehens basierend auf der relativen Zusammensetzung und der Abbaubarkeit einzelner Bestandteile bewertet werden, was für diesen Substanztyp oftmals nicht möglich sein wird. Weiterhin wird eine blockweise Bewertung vorgeschlagen, ohne dass jedoch ausgeführt wird, wie ein derartiger Gruppenansatz praktisch durgeführt werden könnte.

Auch MCS können für die experimentelle Bestimmung der Abbaubarkeit problematisch sein: Einzelbestandteile können zu unterschiedlichem Grade und/oder mit unterschiedlicher Rate abbaubar sein. Im Unterschied zu UVCBs sind MCS aber aus definierten Bestandteilen zusammengesetzt, so dass dies einen Ansatz zur Bewertung der Persistenz ermöglicht, der QSAR und /oder Read-Across zu ähnlichen Substanzen einschließt.

Adsorbierende oder mit der Festphase reagierende Substanzen – Simulationstests mit Sediment und Boden

Wie oben ausgeführt, kann die Interpretation von Simulationsstudien zu Boden und Sediment bedingt durch die Bildung nicht extrahierbarer Rückstände (NER) oftmals eine Herausforderung darstellen. Das Ausmaß der NER-Bildung kann erheblich sein (bis zu 70%, bezogen auf die Massenbilanz des radioaktiven Materials). Im Routineexperiment sind meist keine Aussagen zur Natur der NER möglich. Darüber hinaus fehlt es an einer standardisierten und akzeptierten Methodik für die Extraktion und Charakterisierung gebundenen Materials. Bei der üblichen Quantifizierung von NER durch die isotopenbasierte Massenbilanz können NER sowohl aus xenobiotischen Rückständen bestehen (Mutterverbindung oder Metaboliten) – kovalent und/oder nichtkovalent gebunden an die Matrix – als auch aus biogenen NER, wenn das Radioisotop durch Assimilation in Biomoleküle eingebaut wurde. Während ersteres kritisch zu sehen ist (und entsprechend als nicht abgebaut zu werten), da eine Remobilisierung der Mutterverbindung oder ihrer Metaboliten - selbst im Falle kovalenter Bindung - nicht sicher ausgeschossen werden kann, ist letzteres unbedenklich und als Abbau einzustufen. Allerdings fehlt es gegenwärtig an klaren Definitionen und Methoden, um zwischen diesen beiden Formen differenzieren zu können.

Substanzen hoher Flüchtigkeit

Zwei Methoden stehen für den Test von flüchtigen Substanzen auf biologische Abbaubarkeit zur Verfügung: der "Closed Bottle Test" (OECD 301 D) und der Test nach OECD 310 (CO_2-Headspace test). Für den ersteren ist die Anwendbarkeitsgrenze nicht definiert (keine Obergrenze für die Henry-Konstante angegeben), während mit OECD 310 flüchtige Substanzen bis zu einer Henry-Konstante von 50 Pa m^3 mol^{-1} getestet werden können. Unserem Wissen nach existieren gegenwärtig keine leicht anwendbaren Testsysteme für Substanzen, die eine wesentlich höhere Flüchtigkeit aufweisen.

Forschungs- und Entwicklungsbedarf

Forschungs- und Entwicklungsbedarf wurde hinsichtlich folgender Aspekte festgestellt:

- Verfeinerung und Verbesserung der Standardisierung von Tests zur leichten biologischen Abbaubarkeit bzw. von *Enhanced Screening*-Tests, insbesondere hinsichtlich Inokulumparametern und Validitätskriterien.
- Etablierung verschiedener Sets an Referenzsubstanzen: für schwer wasserlösliche Verbindungen; zum Vergleich bzw. Validierung von *Enhanced Screening*-Tests und kompartimentspezifischen Screening-Tests. Dies ist allgemein von hoher Wichtigkeit, um neuentwickelte oder verbesserte Tests verlässlich bewerten zu können.
- Verbesserung der Leistungsfähigkeit von Screening-Tests hinsichtlich inhibitorischer Substanzen oder Substanzen geringer Wasserlöslichkeit.

- Standardisierung und Ringtests für kompartimentspezifische Screening-Tests als auch für Tests auf inhärente Abbaubarkeit basierend auf CO_2-Bildung.
- Kritische Überprüfung der gegenwärtig angewendeten Prüfkriterien für die Persistenzbeurteilung basierend auf Tests zur inhärenten Abbaubarkeit (OECD 302 B und C) nach ECHA Leitfadendokument R.11: Diese Kriterien waren ursprünglich entwickelt worden, um Halbwertszeiten für Kläranalgen und Umweltkompartimente aus diesen Tests abzuleiten. Dabei sollte insbesondere die mögliche Einbeziehung der CO_2-Entwicklung als zusätzlicher Messgröße berücksichtigt werden.
- Wie zuvor für die einzelnen Tests ausgeführt, bedarf es sowohl einer Bestätigung der Prüfkriterien für *Enhanced Screening*-Tests als auch eines Vorhersagemodells zur Ableitung von P / nicht P aus den Ergebnissen kompartimentspezifischer Screening-Tests.
- Validierung der vielversprechendsten Ansätze für *Enhanced Screening*-Tests, indem diese auf ein Set von Referenzsubstanzen angewendet und die Resultate verglichen werden mit jenen aus Tests auf leichte biologische Abbaubarkeit sowie Simulationstests.

Summary

For evaluating persistence in the context of a PBT/vPvB assessment under REACH currently either screening tests for ready biodegradability (tests of the OECD 301 series) or simulation tests (OECD 307, 308, or 309) are applied. Simulation tests are carried out under environmentally relevant conditions, but are expensive and results depend on many variables and thus are often difficult to interpret. Within the scope of REACH, screening tests for ready biodegradability are mandatory already at the lowest tonnage level. While these tests are characterised by artificial test conditions, not comparable to environmental compartments, these conditions are stringent and interpretation is conservative. Therefore, results from screening tests for ready biodegradability are accepted as screening information for PBT/vPvB assessments.

Due to their conservative nature, many substances require further investigations, when been tested negative in screening tests. Therefore, efforts are made to further develop currently available tests or to develop new ones, which provide reliable conclusions on persistence without the difficulties associated with simulation tests. In the focus of these activities are the so-called enhanced screening tests.

Ideally, (new) screening type tests

- should be significantly faster and cheaper than simulation tests
- should be reproducible and reliable
- should produce reliable outcome with regard to the diverse conditions encountered in the environment (no false negatives for P; a low number of false positives for P)
- should enable predictions for various compartments.

This literature study aims at providing an overview on current developments in the area of biodegradability testing for persistence assessment. Screening tests for ready or inherent biodegradability, enhanced screening tests for ready biodegradability, compartment-specific screening tests and simulation tests are analysed for their advantages and disadvantages with regard to assessing persistence (section 3). Regulatory requirements for identification of persistent compounds are the basis for weighing relevance and applicability of these methodological approaches and are therefore summarized in section 4. Section 5 addresses identified deficits and possibilities for improvement of the five categories of tests with regard to P-assessment. Conclusions on the suitability of available or proposed tests for P assessment are summarised in section 6.

Screening tests for ready biodegradability: deficits and possibilities to improve test performance

The development of the standardised ready biodegradation tests was initiated in the 1970s by different laboratories and each test represents an own history. The different methods have been adopted by the OECD to the OECD 301 A-F guidelines, which have later been complemented by further OECD guidelines (OECD 310) and ISO standards. Chemicals that fulfil the criteria for ready biodegradability in these tests are considered to undergo rapid degradation in the environment under most conditions. A positive result in a ready type test can be assumed as criterion for non-persistency, when the pass levels of 60% ThOD/$ThCO_2$ or 70% DOC-elimination are reached within 28 days, irrespective of the fulfilment of the 10-day window. The tests methods have not been designed to derive degradation rate constants, but merely to measure the removal efficiencies.

In the analysis of existing testing approaches the following deficits in terms of accuracy, repeatability, and comparability have been identified for the ready type screening tests:

- The OECD 301 allows the use of several types of (mixed) inocula, the maximum inoculum density being limited by the validity criterion established for the inoculum blank (e.g. OECD 301 B max.

40 mg /L CO_2 within 28 days). Laboratories could improve the potency of their tests by using inoculum densities which come up near to the upper limit allowed for the inoculum blanks or by using increased flask volumes without questioning the ready type test approach.

- The inoculum concentration of the OECD 301 ready biodegradability tests is mainly limited by the expected values of the inoculum blanks (validity criteria for the blank values), but the existing borderlines of different tests are not consistent. A systematic definition of upper inoculum blank values is lacking. The inoculum used in biodegradation tests should better be described e.g. in terms of the so-called mass loading rate (measured as the biological oxygen demand per total suspended solids per day) of the activated sludge used or additional microbiologic parameters. A better characterisation of the inoculum could help to define the necessary amount of inoculum and lead to a higher reproducibility of tests.
- Attempts to reduce the inoculum blank by pre-incubation and/or pre-treatment may allow higher inoculum concentrations but also have their drawbacks, because often the inoculum potency is reduced. Currently there are no promising routes to further standardise the inoculum (e.g. in terms of defined bacterial strains).
- Many of the studies performed to demonstrate the influence on test results of the inoculum source, quality, pre-treatment, concentration, and total amount have been carried out under test conditions not comparable to standard OECD 301 tests. Thus, the consequences for performing ready type (or enhanced) screening tests still remain unclear.
- The number of replicate vessels in OECD 301 tests is considered as being too low and should be increased, including inoculum blanks (e.g. the MITI-I test requires only one inoculum blank replicate). Further on, criteria for the variability of the inoculum blank values in parallel vessels should be established. The current OECD 301 methods only prescribe a maximum allowed variability of 20% of the degradation extent in parallel vessels, but do not indicate an allowed variance of the inoculum blanks.
- Potential adsorption to activated sludge should be carefully examined when using DOC based tests such as OECD 301A. No guidance on what adsorption extent may be acceptable exists. The elimination through adsorption should be addressed and limited by defining a clear criterion for maximum permissible elimination through adsorption and by including abiotic controls.
- The water-only test systems do insufficiently consider processes like sorption and desorption which may affect bioavailability and degradation.
- For poorly water soluble substances reference substances (substances of poor water solubility and known biodegradability) should be established; this would improve the reliability and comparability of test outcomes for these difficult group of substances. Also, the applicability domain of tests with respect to volatile substances needs to be clearly defined.

These improvements are expected to increase the reproducibility, reliability and comparability (between different tests types, but also between laboratories carrying out the same test), and, by increasing the sensitivity (i.e. reducing the percentage of falsely not identified as biodegradable), improve the applicability of the ready tests in the assessment of persistence.

Enhanced biodegradation screening tests: possible modifications and their suitability for P-assessment

Enhanced screening tests are not designed for determining ready biodegradability of a test substance but exclusively to allow an evaluation as not being persistent. The intention is to avoid the performance of extensive simulation tests. Among the currently discussed modifications (compared to ready biodegradability tests, as the basis for enhanced type tests) are the prolongation of the test duration beyond 28 days, the use of larger test vessels, an increase of biomass concentration, and pre-exposure to the test substance at low concentrations. The inoculum source and quality is considered

being of major importance (like for ready biodegradability tests). The increase of the test vessels and/or the inoculum concentration increases the probability of competent degraders being available. According to REACH guidance document R.7B only test results fulfilling the pass levels of 60% ThOD /$ThCO_2$ or 70% without the application of the 10-day-window should be used to conclude on a substance as not being persistent. Again, the derivation of kinetic degradation rates is not intended because their transferability to natural compartments is not given.

According to REACH Annex XIII, information to be used for persistence assessment includes screening type information with explicit reference to "enhanced ready test(s)" as well as other information, provided that data are suitable and reliable. REACH guidance document R.7B demands that only inoculum ***not*** derived from STP should be used for enhanced biodegradation screening tests to demonstrate non-persistence. As most ready-type tests are performed with STP-derived inocula, and one of the most often applied enhancements is a prolongation of incubation time beyond 28 days, this restriction would significantly limit applicability of this test type. Disqualifying the use of STP derived inoculum is not scientifically justified, as these inocula are allowed for ready type biodegradability tests and these tests are accepted for the P-assessment. Further, inocula derived from STP are introduced in environmental waters via STP effluents (including suspended solids) and generally provide a higher reproducibility compared to, for example, surface water or filtered soil eluates, which highly depend on local or seasonal variations. We therefore conclude that inocula derived from STP should be considered being acceptable for enhanced ready biodegradability testing as far as no pre-adaptation to the test item is envisaged. It must be stressed here, that neither ready biodegradability tests nor enhanced ready tests claim to represent real environmental conditions but aim to provide a generally applicable result on the biodegradability of a compound.

A further obstacle for the use of enhanced biodegradation screening tests in P-assessment is insufficient guidance with regard to possible enhancements and their applicable upper limits as well as possible combinations of enhancements. Thus, there is a need for critical assessments of what modifications are considered being acceptable to yield suitable and reliable information needed for a screening assessment of persistence according to Annex XIII. In our analysis the following issues were identified:

- In REACH guidance on PBT/vPvB assessment the term "enhanced screening test" is not clearly defined although sometimes the term "enhanced ready biodegradation test "is used (REACH guidance document R.11). Thus, it should be clearly stated that enhanced screening tests are restricted to ready type tests and "enhanced inherent screening tests" do not exist. As a consequence the maximum inoculum concentration allowed in enhanced screening test should be below that used for inherent type tests in order to maintain their nature of being ready biodegradability tests.
- The use of inocula from contaminated sites or pre-exposed to low test item concentrations is considered not being suitable for a prospective persistency assessment and should therefore be excluded.
- The prolongation of the test duration (up to 60 days) is not considered being critical because these modifications are already mentioned in OECD and REACH guidance (in case the plateau phase has not been reached). Positive results obtained from prolonged tests can only be used for P assessment (no conclusion on ready biodegradability possible).
- The increase of the test vessels is not considered being critical, because the OECD 301 only gives an indication of a suitable size. The upper limit of the test vessel size will be limited by practical constraints.
- The consideration of *co-metabolism based degradation* is currently not considered in the design of enhanced screening tests. While being an important mechanism of degradation for certain compounds, currently no promising approaches are available in this regard due to the considerable

uncertainty associated with introduction of other carbon sources in tests based on unspecific sum parameters such as DOC-elimination, oxygen depletion, or CO_2 evolution.

- When performing enhanced screening tests, use of additional positive and negative reference substances should be considered, which should reflect the complexity of the test system (i.e. positive reference compounds biodegradable under enhanced but not under ready type conditions as well as negative controls not biodegradable under enhanced conditions).
- The pass levels for enhanced screening tests to be used for persistency evaluation have still not been defined. As a first approach the same pass levels for P/not P equal to that specified within the OECD 301 series could be applied.
- No validity criteria specific for enhanced biodegradation screening tests exist. The criteria given in OECD 301 often are not applicable when certain enhancements are introduced. Differences between measurements of the test and blank vessels should allow accurate estimates of the degradation extent. Use of additional replicate vessels for both the test and blank vessels allow for a more precise estimate of the variability of the measurements and therefore of the accuracy of the biodegradation extent.

The question remains, which enhancement as such and what combinations of enhancements would still be acceptable from the regulatory perspective leading to results which allow conclusions on degradation behaviour (and thereof persistency) under environmental conditions. As a general principle, all combinations of optimizations of ready biodegradation tests which are not questioning their category as "ready type test" should be allowed. Examples are the increase of flask volume, the use of mixed inoculum within the validity criteria, the use of additional positive and negative controls or the increase of replicate vessels. However, the combination of enhancements beyond the test design of ready biodegradability tests, such as an increase of the inoculum density, needs to be critically assessed. Although screening tests are not intended to represent environmental conditions, the use of results from such test systems in the environmental risk assessment should require that the potential for extrapolating from artificial test conditions to environmental conditions is appropriately validated when test designs are modified or enhanced. It is suggested here

1. to define clear limits of single enhancements (as discussed above, e.g. inoculum concentration below that of inherent tests, no acceptance of pre-exposed inoculum, maximum test duration 60 days, test vessel size as appropriate),
2. to evaluate the impact of combining several test modifications and/or enhancements and to exclude combinations which impair the "screening" nature of the enhanced tests,
3. and to critically discuss performance of such enhanced ready tests based on results for substances with known biodegradability. Based on results with these reference substances agreement on suitable pass levels and validity criteria should be achieved.

Inherent type screening tests: deficits, possible improvements and applicability for P-assessment

Inherent tests are performed under more favourable conditions and thus give useful information whether any potential for biodegradation exists irrespective of their relevance for environmental compartments. Results from inherent biodegradation tests may be used for assessing persistency in two ways. First, test results above 70% are used for indicating ultimately biodegradability and are used as trigger for non-persistency when specific criteria (log phase no longer than 3 days, pass level reached within 7 days) are met. Second, negative results from inherent tests (<20% DOC-elimination) indicate a high probability for environmental persistence.

The following problems and possibilities for improvements with inherent tests were identified:

- There is no standardized inherent CO_2 evolution test existing, although in literature several methods based on the OECD 301B and the OECD 310 have been developed The combination of two

endpoints, CO_2 evolution and DOC-elimination in one test has successfully been applied and would more precisely allow distinguishing between mineralisation and adsorption.
- The MITI II test (OECD 302C) requires to use a very specific pre-incubated inoculum mixture, but other inoculum sources such as activated sludge or mixed inocula derived from different environmental compartments seem to be equally acceptable.
- Potential adsorption to activated sludge should be carefully examined when using DOC based tests such as OECD 302B. The three hours value in the Zahn-Wellens test should be indicated in all test reports - not only when there is suspicion for adsorption.

However, the applicability of inherent biodegradability tests for persistency evaluation is restricted because only tests fulfilling the specific criteria described in chapter 4.2.2 allow an assessment as not being persistent. This seems to be justified because of the high inoculum concentration used for these tests, being representative for STPs, only. At the same time, this implies that the upper limit of the inoculum concentration applicable for enhanced screening tests should be well below that used for inherent type tests (e.g. < 200 mg/L d.s. activated sludge).

Despite these potential improvements it remains clear that the inherent tests are not an essential part of a generic strategy to assess persistence, but only may add relevant information in specific cases.

Media-specific screening tests: test designs, deficits and applicability for P-assessment

Media-specific screening tests comprise processes which may affect bioavailability and degradation (e.g. sorption) and thus can provide compartment-specific information on degradation. Such tests allowing for deriving environmentally relevant half-lives (apparent first order) would be closest to simulation tests with regard to media-specific assessment of persistence. However, the high test item concentrations that are needed for achieving good signal-to-noise ratios without using radiolabelled test items hamper the determination of environmentally relevant kinetics from those tests.

Some published but non-standardized media-specific screening tests were developed. It can be assumed that these tests will yield results closer to reality for those media (suspended matter, sediment, soil) as compared to results of screening tests using sewage treatment plant derived inocula without further addition of solids. Namely,

- adsorption processes are mimicked, at least to a certain extent, depending on the exact setup of the test;
- correspondingly, degradation rates are expected to be lower in cases where adsorption is an issue;
- degradation rates will be more media-specific if a media-specific inoculum is used (which may or may not be the case, depending on the exact test system).

To achieve acceptability of those tests for P assessment it may therefore be necessary

- to adapt maximum admissible incubation times for reaching respective pass levels for O_2 consumption or CO_2 production to values above 28 days;
- to adapt valid pass levels for O_2 consumption or CO_2 production to values below the currently foreseen 60%-level based on first results with compounds being well characterized as ready biodegradable in OECD 301-type tests (e.g. aniline).

Up to now no in-depth evaluation has been undertaken and no guidance exists within the scope of REACH regarding these relatively new test types. In principle, media-specific screening tests (e.g. for sediment and soil) can be used as screening tests to indicate rapid biodegradability in the same way

as tests for ready biodegradability, but few experiences exist yet and they still await general acceptance. For regulatory purposes the following criteria should be fulfilled:

4. A consolidated test design should be agreed upon that provides reproducible and sound biodegradation data, suitable to identify (potentially) persistent compounds. This could be achieved by investigating the experimental method in an inter-laboratory comparison (ring test).
5. Taking into consideration the screening type nature of these tests clear criteria need to be developed, which lead to decisions of being (potentially) persistent or not. Such prediction models might use degradation pass levels with or without specific degradation rates or other test outcomes but will ultimately lead to a yes/no conclusion on persistence. In general, direct comparison with Annex XIII P-criteria is not recommended, as test conditions do not reflect environmental conditions.
6. The resulting test designs and criteria then should be applied to determine biodegradation data for a defined set of test compounds of known and differing degree of biodegradability to check modified and new biodegradability test methods.
7. The results should finally be compared to results for the same compounds obtained from standardized screening test according to OECD 301 and – if available – to biodegradation data from literature and databases from well-established simulation studies for the respective compartment.

In conclusion, newly developed compartment-specific tests (for soil and sediment) are considered screening tests by nature. They may be promising tools for obtaining compartment-specific information on degradation, in cases where there are indications that a substance is mainly distributed to that specific compartment or that degradation in that compartment is different than in others (e.g. pelagic compartment). Their precise role within a testing strategy for persistence assessment needs to be established.

Simulation tests on degradation in environmental media: deficits and possible improvements

Several deficits and gaps of existing simulation tests have been identified and discussed over the past years. The main issue with regard to OECD 307 on "Aerobic and Anaerobic Transformation in Soil" relates to the question how non-extractable residues (NER) and bound residues (BR) should be defined, determined and interpreted.

Main points of discussion and criticism in relation to OECD 308 (water sediment simulation) can be summarized as follows:

- The continuous exchange between water and sediment in the test system is neither standardized nor quantifiable. Therefore, it is not possible to determine robust compartment-specific degradation half-lives.
- The redox conditions in the sediment are undefined, since the test simulates an aerobic water column over a thin aerobic sediment layer followed by a deeper anaerobic layer. Consequently, the test results are influenced by factors that affect the oxygen distribution in the sediment, e.g. aeration rate, turbulence in the water phase, sediment depth and texture.
- The recommended water:sediment ratio of 3:1 to 4:1 (v/v) together with the recommended sediment height of 2.5 ± 0.5 cm does not represent 'natural' conditions. Due to the small ratio equilibrium mass distribution is often shifted towards the sediment phase.
- The resulting range of system geometries (e.g. height of water and sediment layer, interfacial area) can influence distribution processes (being governed by the partitioning equilibrium constant and the diffusion rate from the water/sediment interface into bulk sediment) as well as biodegradation processes and thus affect persistence in the experimental system. Hence, results are to some extent an artefact of the test system, which hampers the transferability to environmental conditions

- The static design of the test also does not represent natural conditions by not accounting for the effects of flow velocity and sediment dynamics.
- The total system half-life (water-sediment) obtained from OECD 308 studies are difficult to compare with Annex XIII criteria for individual compartments.

Several possibilities to improve test performance and evaluation of OECD 309 studies have been made recently:

- The wide array of experimental options in OECD 309 leads to very different outcomes. The test setup should therefore be specified (e.g. with regard to the appropriate temperature to be used (e.g. 12 or 20°C). Test systems should be shaken rather than stirred to prevent increased sorption during the experiment due to sediment grinding
- The option to include "diffuse light" must be elaborated: clear specifications on light-dark cycles should be given; intensity, duration and wave length distribution must be defined.
- New modelling approaches might allow the prediction of compartment-independent biodegradation parameters (e.g. the bioavailability-corrected and biomass-normalized second-order biodegradation rate constant k'_{bio}) that can be used to assess persistence at the water-sediment interface (also relevant for OECD 308).

In conclusion, simulation tests are currently the last resort for biodegradability testing, because they are time- and resource intensive. Although they haven't been in the focus of this review possibilities for improvements were identified, which can be summarised as follows:

3. The OECD 307 test is well established for testing biodegradability in the soil department. As for any soil test the identification and evaluation of non-extractable residues is an issue that is not yet finally resolved.
4. OECD 308 and 309 would benefit from a specification of test designs and test performance criteria.

Poorly water soluble substances

Currently, poorly water soluble substances[2] in principle can be assessed for degradability using screening type tests, especially if specific modifications are introduced to account for this difficult substance class. However, compared to soluble compounds there is an imminent risk of underestimating biodegradability as the effectively available concentration may be low in screening type tests (limited by absolute solubility, kinetic constraints on solubilisation or the accessible surface area). As a consequence, in extreme cases there may be no relevant growth of biomass. As biomass growth determines the observed sigmoidal biodegradation curve typical for these tests, biodegradation may be severely underestimated. In contrast, due to the much lower concentration present in simulation tests, the nominal concentration applied may be close or below the solubility limit even for poorly soluble compounds and effective degradation may be observed (if the compound is not per se recalcitrant to biodegradation).

In order to improve performance and interpretation of (enhanced) ready tests for this group of substances (additional) reference compounds of poor solubility (e.g. microcrystalline cellulose) but sufficient biodegradability should be validated and used.

---

[2] There is no uniform definition of poor water solubility available: REACH guidance R.7b refers to OECD (< 100 mg/L) and EU-TGD (2003) (< 1 mg/L), outlining that rather below 1 mg/L problems are expected to occur with regard to environmental fate and ecotoxicity testing,

UVCB and multi-constituent substances

A systematic and practically feasible concept for the assessment of *Substances of Unknown or Variable Composition, Complex Reaction Products and Biological Materials* (UVCB) is lacking: while an assessment of single constituents down to 0.1% w/w is demanded within the scope of REACH PBT/vPvB assessment, individual constituents of UVCBs are required to be specified down to concentrations of ≥ 10% only. Other constituents are often poorly defined, rendering radioactive labelling practically impossible, as well as application of QSAR models. Thus, while screening tests may be the only practicable way of demonstrating degradability, REACH guidance restricts their use to the case of UVCB consisting of homologous structures. In case of UVCBs based on structurally non-homologous constituents, biodegradability "should be judged on a case-by-case basis" based on "relative composition and degradability of individual constituents" which often may not be possible for this type of substance. Further, a "block-wise" evaluation is suggested for these UVCBs, but no detailed guidance is given on how to carry out such a grouping approach.

Also multi-constituent substances (MCS) might pose problems for experimental testing of biodegradability: Individual constituents may degrade to a different extent and / or at a different rate. But different to UVCBs, constituents of MCS are defined and allow QSAR and/or read-across approaches to characterise biodegradability.

Substances adsorbing to or reacting with matrices - simulation tests with sediment and soil

As outlined above, interpretation of simulation studies in soil and sediment is often challenging due to the formation of non-extractable residues (NER). The extent of NER formation can be remarkable (up to 70% with regard to mass balance of radioactive material). The nature of NER is mostly not assessed by routine experiments. Moreover, methodology for extraction and characterization of bound material is far from being standardised or generally agreed. Based on isotope mass balancing, NER may be composed of xenobiotic NER (parent or primary metabolites) bound to the soil matrix either by non-covalent or covalent interaction; or biogenic NER, where the isotope label is assimilated in biomolecules. While the latter is regarded to be of no concern, both, parent or primary metabolites may be remobilized even when covalently bound and therefore may not be regarded as degraded. Clear definitions and methods to differentiate between these forms are lacking.

Substances of high volatility

Volatile substances can be tested according to two methods: using the closed bottle test (OECD 301 D) (the applicability domain is not defined: no upper limit for the Henry coefficient is given) or according to OECD 310 (CO_2-Headspace test, for substances with Henry coefficients up to 50 Pa m^3 mol^{-1}). To our knowledge, there are currently no readily applicable test systems allowing for testing of compounds of pronouncedly higher volatility.

Needs for further research and development

The following research and development needs were identified:

- Refinement and improvement of standardisation of tests for (enhanced) ready biodegradability, especially with regard to inoculum parameters and validity criteria.
- Establishment of sets of reference substances (for poorly water soluble substances, for comparing results from enhanced tests and from compartment-specific tests) is important to reliably assess newly developed or amended tests.
- Improvement of screening test performance for inhibitory or low water soluble substances.
- Standardization and ring tests for compartment specific screening tests as well as for inherent type tests with CO_2 evolution.
- Critical review of the currently applied pass levels for P with regard to inherent tests (302B and C; especially when combined with CO_2 evolution) as given by ECHA guidance document R.11, which were actually developed as a prerequisite to derive half-lives for STP and environmental compartments.
- As detailed above for the particular test types, pass levels for enhanced screening tests need confirmation; and a prediction model with regard to P / not-P needs to be established for compartment-specific screening tests.
- Validation of the most promising developments with enhanced screening tests by applying them to a set of reference compounds and comparing them to test results from tests for ready biodegradability and simulation tests.

1 Introduction

For assessment of persistence in the context of a PBT/vPvB assessment under REACH currently either screening tests for ready biodegradability (tests of the OECD 301 series) or simulation tests (OECD 307, 308, or 309) are applied. Screening tests for ready biodegradability are easy to carry out and well established, but use artificial test conditions, not comparable to conditions in the environmental compartments, for which persistence of a substance should be assessed. Screening tests are conservative (i.e. leading to few non-biodegradable substances identified as biodegradable, "false positive results") and only allow yes/no conclusions regarding persistence. In consequence, many substances require further investigations, when been tested negative in screening tests. On the other hand, simulation tests are carried out under environmentally relevant conditions, but are expensive and dependent on many variables. Therefore, efforts are made to further develop currently available tests or to develop new ones, which provide reliable conclusions on persistence without the difficulties associated with simulation tests. In the focus of these activities are the so-called enhanced screening tests.

This literature study aims at providing an overview on current developments in the area of biodegradability testing for persistence assessment. Screening tests for ready or inherent biodegradability, enhanced screening tests for ready biodegradability, compartment-specific screening tests and simulation tests are analysed for their advantages and disadvantages with regard to assessing persistence. Based on this analysis test modifications discussed in the literature are critically assessed and recommendations are given how to improve test designs.

2 Data sources and methodology

The OECD testing guidelines for testing of chemicals (OECD, 2016) provide the basis for the performance and interpretation of standardized biodegradability tests. The OECD guidelines 301 A-F as well as OECD 310 describe different methods for determining ready biodegradability, OECD 302 A-C methods for inherent biodegradation. These screening tests are complemented by different simulation tests such as OECD 303, OECD 307, OECD 308, OECD 309 or OECD 314. In parallel, several biodegradability standards have been developed by ISO TC 147, which mostly are related or preceding to the respective OECD methods. Further standards developed by ASTM of the US EPA have not been considered systematically within this study.

In the OECD Series on Testing and Assessment several review documents on the assessment of biodegradability and persistence have been published. In 1995 Painter et al. elaborated a review paper on biodegradability on behalf of the OECD (OECD, 1995). In 2004 a workshop of Simulation Testing of Environmental Persistence was organised by the Netherlands Organisation for Applied Scientific Research (TNO) on behalf of the Dutch Institute for Public Health and the Environment (Bowmer and Leopold, 2004). Participants for the first time indicate the need to develop a new generation of tests to fill the gap between screening tests and simulation tests using radiolabelled substances. As a result of this workshop further initiatives on improvements of biodegradability and persistency assessment have been undertaken by the European Centre for Ecotoxicology and Toxicology of Chemicals (ECETOC) and the long research initiative of the European Chemicals Industry Council (CEFIC) (ECETOC, 2003; 2007; 2011; 2013c; 2014). These reports provide the basis for the literature research.

Database searches were performed applying a tiered approach:

- searches in Pubmed (http://www.ncbi.nlm.nih.gov/pubmed/) were used to identify key issues, main search terms, the amount of information to be expected as well as to identify a large part of the relevant literature for the search period 2005-2015; hits were screened for relevance using titles and abstracts;
- the search strategy was applied to the databases Science Citation Index and Biosis Previews at host DIMDI, Cologne (http://www.dimdi.de/) and extended to the period >2000; hits were compared to previous findings by titles and relevant new articles were retrieved;
- these searches were supplemented by specific searches in contents of journals: Science Direct (Elsevier, e.g. Chemosphere), Springer (e.g. Environmental Science and Pollution Research), and SETAC (e.g. Environmental Toxicology and Chemistry) have been systematically evaluated.

The literature research was complemented by author-related searches for publications from well-known researchers such as Battersby, Boethling, Ericson, Nyholm, Painter, Parsons, Seyfried, Thouand etc. as well as via a Google search for grey literature of the last 5 years.

3 Current testing approaches

3.1 Screening tests for ready biodegradability

3.1.1 Overview

The development of the standardised ready biodegradation tests was initiated in the 1970s by different laboratories and each test represents an own history. The different methods have been adopted by the OECD to the OECD 301 A-F guidelines, which have later been extended by further OECD guidelines (OECD, 2016) and ISO standards. Chemicals that fulfil the criteria for ready biodegradability are considered to undergo rapid degradation in the environment under most conditions, provided that testing follows the stringent conditions of the OECD 301 methods (ECHA, 2014b). The tests methods have not been designed to derive kinetic degradation rate constants, but merely to measure the removal efficiencies (Kaiser, 1998). The objective of ready biodegradability testing is to predict whether a chemical will degrade in specific environmental situations, but not the extent of biodegradability in these situations (OECD, 1995).

The principle of standard ready biodegradability tests is the incubation of the test and reference substance as the only organic carbon source with an inoculated mineral medium in the dark or diffuse light at 22 ± 2 °C for 28 days. Degradation is followed by DOC analysis, or measuring CO_2 evolution or oxygen consumption at frequent intervals. The activity of the inoculum alone is considered in parallel blank control flasks.

The OECD 301 guidelines (OECD, 2016), which have also been included into Regulation (EC) No 440/2008 laying down test methods pursuant to REACH, have the following main characteristics described in Figure 1 (data from OECD 301, 1992).

Figure 1: Main characteristics of ready biodegradation tests

		Endpoint	Appropriate for substances			test conc.	Inoculum	Inoculum	Potency *)
			low soluble	volatile	adsorbable	mg/l		10^4 cells / L	
OECD 301 A	DOC Die-Away	DOC	-	-	+/-	10 -40 DOC	30 mg d.s/L 100 ml/L	1000-10000	++
OECD 301 B	CO2 Evolution	CO_2	+	-	+	10 -20 TOC	30 mg d.s/L	1000-10000	++++
OECD 301 C	MITI (I)	O_2 / DOC	+	+/-	+	100	30 mg d.s/L	1000-10000	++
OECD 301 D	Closed Bottle	O_2	+/-	+	+	2 - 10	0.05 - 5 ml/L	1-100	+
OECD 301 E	Modified OECD Screening	DOC	-	-	+/-	10 -40 DOC	0,5 ml/L	10	+
OECD 301 F	Manometric Respirometry	O_2	+	+/-	+	50 - 100 ThOD	30 mg d.s/L	1000-10000	+++
Comparable methods									
OECD 306	Biodegradability in Seawater	O_2	+/-	+	+	2	100% seawater		+/-
		DOC	-	-	-	5-40 TOC			+/-
OECD 310	CO_2-Headspace test	CO_2	+/-	+/-	+	2-40 TOC (20 mg TOC)	4 - (30) mg d.s/L	100-1000	+++
ISO 10708	Two-phase closed bottle test	O_2	+/-	+/-	+	100 ThOD	30 mg d.s/L		++

DOC: Dissolved organic carbon, IC: inorganic carbon

*) Experience of routine GLP laboratory Hydrotox

Figure 1 demonstrates that the different test methods have their limits regarding the applicability to test substances and differ from one another in their test concentration and inoculum density and the biodegradation potential (potency).

Due to the respective development history of the tests there are differences, which cannot be explained by different conditions or intentions. For example, the CO_2 headspace test (OECD 310) allows test concentrations between 2 and 40 mg C/L while the CO_2 evolution test (OECD 301 B) only allows test concentrations between 10 and 20 mg C/L. Both tests may therefore differ considerably in their inoculum-substrate ratio, since also the amount of activated sludge inoculum differs considerably. There is, however, no scientific reason why the CO_2 evolution test may not be tested at higher test concentrations up to 40 mg/L or why the inoculum concentrations may not be reduced to 4 mg d.s./L in order to allow lower test concentrations below 10 mg/L.

The assessment of biodegradation in the marine environment is a special case for which OECD 306 has been developed. In general, the expected biodegradation in seawater tests is lower than in fresh water tests, because of higher salt concentration and lower temperature, both resulting in lower microbiologic activity. The guideline describes adaptations for the OECD Screening test (shake flask test, OECD 301 E) and the Closed bottle test (OECD 301 D) with respect to the inoculum (100% sea water), the incubation temperature (15-20°C), and the test duration (up to 60 days). The tests results are not taken as indicator for ready biodegradability, but for obtaining information about the degradability in marine environment. In contrast, REACH guidance R 7b (ECHA, 2014b) states that positive results of the OECD 306 test are a strong indicator that the criteria for ready biodegradability are also fulfilled.

ISO 16221[3] provides further guidance for biodegradability testing in the marine environment. Here, next to the DOC shake flask test and the Closed bottle tests, also the two-phase closed bottle test (ISO 10708**)** as well as the CO_2 evolution test and the CO_2 headspace test have been adapted for marine conditions. Natural seawater or artificial seawater may be used. Artificial seawater is inoculated with marine seawater, marine suspended sediments or with bacteria from marine aquariums. The test vessels are incubated within the range of 15°C to 25°C (± 1°C) for up to 60 days.

3.1.2 Inoculum

3.1.2.1 Inoculum source

According to the OECD introduction to ready biodegradability, "the inoculum may be derived from a variety of sources" such as "activated sludge; sewage effluents (unchlorinated), surface waters and soils; or from a mixture of these." When activated sludge is used, it should be taken from a treatment plant or laboratory-scale unit receiving predominantly domestic sewage. Inocula from other sources, usually yielding lower cell densities, have been found to give higher scattering of results (OECD 301, adopted 1992, paragraph 17; OECD, 2016).

According to OECD (OECD/OCDE, 2006) "it has been recognised that standardisation of the inoculum might also improve the comparability of the methods. However, it was concluded that this is not possible without significantly reducing, at the same time, the number of species present in the test system. A mixed inoculum is therefore recommended to ensure the presence of a variety of degrading organisms in the tests. In view of the stringent requirements to these tests, it was also decided that pre-exposure (i.e. pre-adaptation) of the inoculum to the test substance should not be allowed. If pre-exposed inoculum was used, the test is per definition no longer a test for ready biodegradability, and a positive result may then be used to classify the test substance as 'inherently biodegradable with pre-adaptation'".

[3] ISO 16221:2001 Water quality -- Guidance for determination of biodegradability in the marine environment.

3.1.2.2 Limitation of the biomass

According to OECD 301 (adopted 1992; OECD, 2016), section 4, the amount of DOC introduced with the inoculum should be kept as low as possible when compared to the organic carbon introduced with the test substance. A number of pre-treatment approaches are usually applied, in order to reduce the level of organic carbon introduced with the inoculum and/or to reduce the overall respirometric activity of the inoculum. The inoculum may be pre-incubated [4] over several days and/or may be washed with tap-water or dilution water by settling/centrifuging and re-suspending. The objective of these pre-treatments always is to reduce the inoculum blank values and thus, to improve the accuracy of the calculation of degradation extents. The validity criteria of the tests often limit the inoculum concentration to be applied in the tests. There are several studies which confirm that usually these pre-treatments reduce the capability of the microorganisms as competent degraders (see section 3.3.2.2). While the pre-incubation to the experimental conditions is considered being acceptable, OECD 301 (adopted 1992, paragraph 18; OECD, 2016) excludes a pre-adaptation to the test substance. Further guidance on any pre-treatment of the inoculum, such as washing, filtration or centrifugation of the inoculum source, is not given. There have been several approaches for a better harmonisation and characterisation of the inoculum used and of the consequences of these pre-treatment steps which are discussed in detail in chapter 3.3.2.2. In principle, many of these proposals could also be applied for standard OECD 301 testing, without questioning the ready type test category.

The implications of the different, and sometimes contradicting, validity criteria with respect to inoculum activity in the blanks are further discussed in chapter 3.1.3.2.

3.1.3 Test design

3.1.3.1 Applicability of test methods

In the OECD 301 introduction to ready biodegradability (adopted 1992, paragraph 9; OECD, 2016) some guidance on the applicability of the different tests is given. Water-soluble, non-volatile and non-adsorbing test substances may be assessed in all tests. For substances with low water solubility and those, which tend to adsorb e.g. to the inoculum, the DOC-die away test (OECD 301A) or the Modified OECD Screening test (OECD 301E) are not the first choice, because these use DOC-measurements as end-point. The CO_2-evolution test (OECD 301B) is continuously aerated with CO_2-free air and is not applicable to volatile substances. Methods, which use a headspace, such as the MITI(I) test (OECD 301C), the Manometric respirometer test (OECD 301 F) or the CO_2-Headspace test (OECD 310) may be applied for moderately volatile substances (see Figure 1).

ISO/TR 15462[5] provides further guidance on the application of ISO biodegradation tests for the aquatic environment. Next to determining the biodegradability of chemical substances these methods are designed for testing environmental samples such as wastewaters.

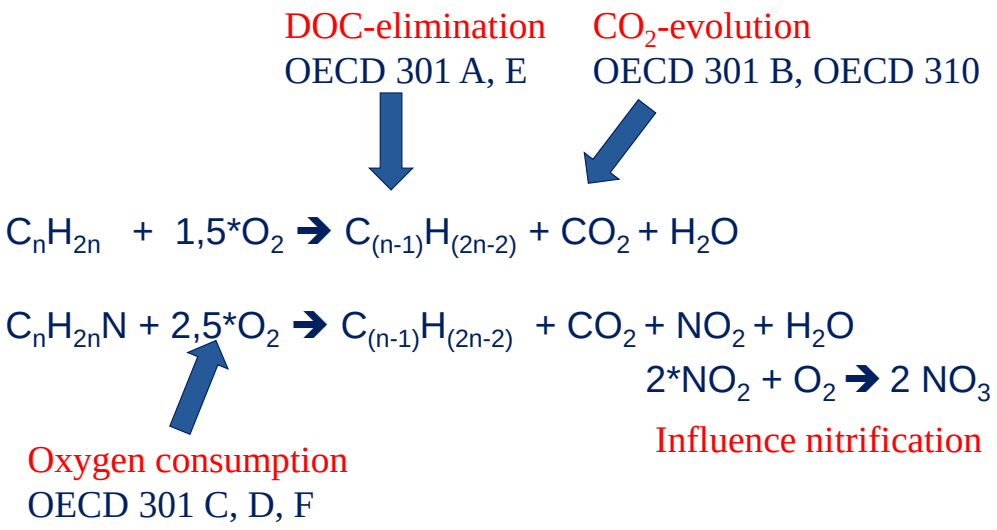
While all tests of the OECD 301 series are considered to describe ultimate biodegradation (mineralisation), the endpoints used are not unambiguous. CO_2-evolution is a definite proof of mineralisation while oxygen consumption is an indirect proof of mineralisation and is additionally influenced by nitrification processes (Figure 2).

---

[4] In this report the term "pre-incubation" is used for the pre-incubation of the inoculum to the test conditions mainly applied to reduce the inoculum blank values, while the term "pre-exposure" is used for any pretreatment of the inoculum in presence of the test substance with the aim to obtain pre-adapted inoculum.

[5] ISO/TR 15462:2006 Water quality -- Selection of tests for biodegradability

Figure 2: End points for assessing biodegradability



The oxygen consumption for nitrification may cause considerable uncertainty in data processing, especially for nitrogen containing test substances. Annex V of the OECD 301 introduction provides a correction method for considering nitrification processes via nitrite and nitrate analysis, which, however, considerably increases the effort for the test performance (OECD 1992). There have been attempts to avoid nitrification in respirometric test systems by adding the nitrification inhibitor Allythiourea (10 mg/L) (Stasinakis et al., 2008). This is not *in compliance with* the OECD 301 F test guideline.

DOC elimination should only be interpreted as mineralisation, if no substantial adsorption and/or volatilisation of the test substance take place.

There have also been attempts to combine several endpoints such as CO_2 evolution and DOC elimination. As regards the carbon dioxide evolution test according to ISO 9439[6], the informative Annex D suggests additional DOC measurements at the start and the end of the test or at regular time intervals. The test extension described in Annex D also allows for higher inoculum concentrations up to 150 mg/L (dry solids activated sludge). In this case the CO_2 evolution of the inoculum blanks should be in the range of 150 mg/L). This test design corresponds more to an inherent test than to a ready test (for comparison: the lowest inoculum concentration of the Zahn-Wellens test is 200 mg /L activated sludge).

3.1.3.2 Replicate vessels and inoculum activity

The OECD test guidelines provide differing recommendations for the inoculum source and density, the volume of the vessels/flasks used and the number of parallel flasks for the test substance and the inoculum blank (Figure 3).

[6] ISO 9439:1999 Water quality -- Evaluation of ultimate aerobic biodegradability of organic compounds in aqueous medium -- Carbon dioxide evolution test.

Figure 3: Test design and specific validity criteria for ready biodegradation tests

Ready biodegradation tests

		Endpoint	Vessel size	Replicates		Validity inoculum blank	
			mL	test	blank	mg/L O_2 or CO_2	
						range	max.
OECD 301 A	DOC Die-Away	DOC	e.g. 250-2000	2	2	none	
OECD 301 B	CO2 Evolution	CO_2	2000-5000	2	2	40	70
OECD 301 C	MITI (I)	O_2 / DOC	300	3	1	20-30	60
OECD 301 D	Closed Bottle	O_2	100-300	10	10		1,5
OECD 301 E	Modified OECD Screening	DOC	250-2000	2	2	none	
OECD 301 F	Manometric Respirometry	O_2	suitable	2	2	20-30	60
Comparable methods							
OECD 306	Biodegradability in Seawater	O_2	300	8	8	< 30% of test bottles	
		DOC	e.g. 250-2000	2	2	none	
OECD 310	CO_2-Headspace test	CO_2	e.g. 160	ca. 20	ca. 20	< 3 mg IC /L	
ISO 10708	Two-phase closed bottle test	O_2	200-300	3	3	< 3 mg/L (1st week), < 1 mg/L per week	

DOC: Dissolved organic carbon, IC: inorganic carbon

The inoculum activity is monitored by testing a reference compound and by determining the oxygen uptake or CO_2 evolution of the inoculum blank vessels. A positive result with a reference substance is a prerequisite for all OECD 301 tests for being valid. Further on, for those tests which use oxygen consumption or CO_2 evolution as endpoints, a maximum allowed inoculum activity has been defined in the validity criteria. No corresponding validity criterion regarding the inoculum activity exists for the OECD tests using DOC measurements. In contrast, ISO 7827 limits the DOC introduced with the inoculum to 10% of that introduced with the test substance (see below).

The reasons for establishing these validity criteria are the following:

- There is an initial maximum oxygen supply of 8-9 mg/L available in the Closed Bottle Test (OECD 301D).
- The inoculum blank values are subtracted from the test substance values. Thus, the accuracy is reduced, if these differences are obtained from high inoculum blank values.
- The validity criteria reflect typical values observed in these tests and exceeding inoculum blank values may indicate failures of the test system such as leakages (e.g. in the CO_2-evolution test OECD 301B).

The strict validity criteria referring to the inoculum blanks may be difficult to fulfil. As a consequence, a pre-incubation or pre-treatment of the inoculum often becomes necessary, which generally lowers the biological potency (see chapter 3.3.2.2).

The CO_2 Headspace test according to ISO 14593 has as a validity criterion that the total inorganic carbon (TIC) produced in the blank controls at the end of the test is less than 15% of the organic carbon introduced as test substance.[7]

The minimum number of replicate test vessels and inoculum blank vessels prescribed in the different test methods is usually only two, with the exception of those tests, where the flasks are sacrificed for the measurement (OECD 301 D and OECD 310).

One option to improve the reliability of biodegradation tests consists in increasing the number of replicate vessels for both the test substance and the inoculum control and in defining maximum deviations between replicate vessels in terms of the standard deviation (see chapter 5).

3.1.3.3 New testing approaches

There have been several proposals to refine the existing screening tests. Some of these proposals aim at prolonging the test duration, increasing biomass or increasing the size of test vessels. These are discussed in detail in chapter 3.3. Other proposals are not contradictory to the general test conditions prescribed for ready biodegradation tests. They comprise e.g. consideration of additional sterile controls without inoculum to recognize relevant hydrolysis, complementary chemical analyses of the parent compound and (known) transformation products in the test vessels, or additional nitrogen analyses for describing nitrification processes. Another proposal is to consider an additional procedure control without inorganic nitrogen in order to identify substances which may be readily biodegradable as nitrogen source, but not degradable in the presence of nitrogen containing test medium (e.g. Wess and Eisner 2014)[8].

Further proposals describe completely new test systems which apply conditions that are different from the OECD test conditions:

A new miniaturised high throughput screening system based on the Closed bottle test was developed recently (Cregut et al., 2014). Oxygen concentration is monitored by a non-invasive fluorescent oxygen opto-sensor dye. The test can be performed in 24 well plates. As the absolute inoculum amounts per well are relatively low due to the small test volume, this could lead to random failures and highly variable lag phases due to an insufficient amount of specific degraders. On the other hand, the possibility of performing multiple tests in parallel with different types of inocula may outweigh this disadvantage and may even provide a more representative pattern for biodegradability.

A thoroughly new test design with the intention to produce an output comparable to OECD 301 tests was suggested recently (Czechowska et al., 2013). Instead of monitoring CO_2 evolution, O_2 consumption or DOC decline, microbial community growth is monitored over time by flow cytometry counting. The biodegradability is measured by cell counting and assuming conversion factors of 0.4 pg C per cell for activated sludge bacteria, and of 0.2 pg C per cell for freshwater lake bacteria. This allows testing at low concentrations (1-2 mg C/L) and obtaining the results within 6-9 days. One disadvantage of the method is that substance dissipation cannot be directly measured and no accurate compound mass balance can be established. The authors tested 2-hydroxybiphenyl next to a set of six fragrance ingredients differing in water solubility and vapour pressure. A reasonable agreement with results of standardized OECD 301 test was shown.

---

[7] ISO 14593:1999 Water quality -- Evaluation of ultimate aerobic biodegradability of organic compounds in aqueous medium -- Method by analysis of inorganic carbon in sealed vessels (CO_2 headspace test)

[8] Wess R.A. Eisner, G. (2014) Refinement of biodegradation tests to prepare for subsequent E-fate testing and assessment. Poster presentation SETAC Europe Basel 2014

It is questionable, whether these new test designs will be acceptable for regulatory purpose in the near future. They may, however, have their advantages for screening purposes within research and development.

3.1.4 Interpretation of results

The revised OECD introduction to the degradation testing of organic chemicals (OECD/OCDE, 2006) outlines that "Ready biodegradability tests must be designed so that positive results are unequivocal." A positive result in a test of ready biodegradability is interpreted in such a way that the chemical will undergo rapid and ultimate biodegradation in the environment. From the regulatory perspective no further investigation of the biodegradability is required for readily biodegradable substances and possible environmental effects of transformation products are not assessed. Because ready biodegradability tests may sometimes fail because of the stringent test conditions, consistent positive test results from test(s) should generally supersede sporadic negative test results. This principle is also applied according to REACH guidance R7b (page 192) (ECHA, 2014b). When conflicting test results are reported, it is recommended to check the origin of the inoculum and whether possible adaptation of the inoculum may be the reason. "According to regulatory authorities a careful assessment of the available results and study descriptions is necessary when positive and negative tests are occurring in comparable test systems.

A negative result in a test for ready biodegradability does not necessarily mean that the chemical will not be degraded under relevant environmental conditions, but it means that the next level of testing, i.e. either a simulation test or an inherent biodegradability test, should be considered".

According to OECD 301 (1992, paragraph 24) a test is considered valid, if the maximum difference of biodegradation in replicate vessels at the plateau, at the end of the test or at the end of the 10-d window, as appropriate, is less than 20% [on the absolute degradation scale from 0-100%]. A second validity criterion is that the reference compound has reached the pass levels by day 14. It is stated that "because of the stringency of the methods, low values do not necessarily mean that the test substance is not biodegradable under environmental conditions, but indicates that more work will be necessary to establish biodegradability."

While the maximum difference of degradation extents in parallel vessels of 20% has been set as validity criterion for the test substance, no such criterion exists for the validity of replicate inoculum control vessels. However, a high variability in the blanks could significantly influence the overall test results, because the mean value of the blanks is subtracted from the distinct test vessel values. With a higher number of replicate vessels (for both the test and inoculum blank vessels) a better description of the variability and the identification of outliers could be achieved (see chapter 5).

The pass levels for ready biodegradability are set to 70% DOC removal and 60% of ThOD or $ThCO_2$ production for respirometric and CO_2 evolution methods. The reason for these different pass levels is that some of the carbon from the test chemical is incorporated into new cells and thus the percentage of CO_2 produced is lower than the percentage of carbon being used. These pass values have to be reached in a 10-d window within the 28-d period of the test, with exceptions as mentioned below. The 10-d window starts when the degree of biodegradation has reached 10% DOC, ThOD or $ThCO_2$ and must end before day 28 of the test is reached. Chemicals which reach the pass levels later than 28-days are not considered to be readily biodegradable. The 10-d window concept is not applied to the MITI method (OECD 301, paragraph 10).

Further guidance concerning the interpretation of DOC-elimination is given in ISO 7827:[9]

- The expression of results is only assigned to biodegradation, if the substance is not significantly removed abiotically (e. g. by adsorption) and the removal curve has a typical shape with a lag and a degradation phase.
- The DOC contribution to the test system by the inoculum should be less than 10% of that of the test compound (validity criterion).

3.1.5 Rules for difficult substances

3.1.5.1 UVCB substances

UVCB substances are multi-component mixtures which comprise different individual substances with different solubilities or physico-chemical properties. In most cases they are characterised by certain ranges of carbon-chain lengths and branches or degrees of substitution and positions. Examples are fatty-acid ethoxylates, quaternary ammonium compounds, or petroleum substances derived from crude oil.

For complex mixtures a stepwise, sequential adaptation of the microorganisms to the individual substances contained is often observed, which leads to different and overlapping biodegradation kinetics (REACH Guidance R7b, page 253). Thus, for UVCB substance the proof of ultimate biodegradability is accepted also if the 10-day-window as criterion for ready biodegradability is not achieved or if the test duration has been extended.

For example, the Regulation (EC) No. 648/2004 on detergents requires ultimate degradability of surfactants in ready type tests without applying the 10-day window principle. Also according to REACH guidance, the 10-day window does not apply for a mixture of homologous compounds (ECHA, 2014b). For substances or plastic materials with low water solubility also longer test durations up to 6 months have been accepted (ISO 14851, ASTM D5864).[10, 11]

The 10-day window as criterion for ready biodegradability has been criticised as a suitable approach for describing biodegradation kinetics of mixtures and poorly water soluble substances, because the different homologs and isomers will be degraded sequentially (Richterich and Steber, 2001) and is not applied for high volume surfactants used in detergents which are released to municipal waste water treatment plants. It is generally irrelevant for persistence assessment with regard to PBT/vPvB properties (ECHA, 2014b; c).

3.1.5.2 Poorly water-soluble substances

The OECD 301 (adopted 1992, paragraph 19; OECD, 2016), refers to different methods for adding the test and reference substances depending on the nature of the chemical. Water soluble substances may be added via stock solutions to the test vessel, whereas substances with low water solubility can be directly added to the final mineral medium. Annex III of the OECD guideline explicitly refers to the

---

[9] ISO 7827:2010 Water quality -- Evaluation of the "ready", "ultimate" aerobic biodegradability of organic compounds in an aqueous medium -- Method by analysis of dissolved organic carbon (DOC).

[10] ISO 14851:1999: Determination of the ultimate aerobic biodegradability of plastic materials in an aqueous medium -- Method by measuring the oxygen demand in a closed respirometer

[11] ASTM D5864-11: Standard Test Method for Determining Aerobic Aquatic Biodegradation of Lubricants or Their Components.

handling of poorly soluble and insoluble substances[12]. It is recommended that solid materials should be homogenised to avoid errors due to non-homogeneity. The use of an emulsifier or solvent, which should not be toxic to bacteria and must not be biodegraded or cause foaming under test conditions, is allowed. In this case additional blanks containing the auxiliary substance (emulsifier or solvent) should be included. Oily substances might also be applied to solid carriers. However, according to the MITI method (301 C) neither organic solvents nor emulsifying agents should be used.

According to the REACH guidance document R.7b (ECHA, 2014b) (Appendix R.7.9-3), a number of modifications of test item addition to the ready biodegradability tests have been suggested, which are based on the OECD 301 and ISO Guidance 10634[13]. Using these adaptions for poorly water-soluble compounds will not preclude a substance from being identified as readily biodegradable.

Overall, these measures are

- Direct addition by weighing, using an inert support such as silica gel or glass filters, where appropriate. Adsorption to an inert support may be realised by means of a carrier solvent, which is later removed by (rotatory) evaporation. Naturally, the method is not applicable to volatile substances. The silica gel method was published (Handley et al., 2002) as a possible "standard method of adding low density, poorly water-soluble substances into test vessels of biodegradability studies to ensure these materials remain in contact with micro-organisms in the test medium." It is suggested by the authors to be environmentally relevant as similar processes were expected to occur in the environment.
- A solution of the test compound is prepared in a volatile organic solvent and is introduced into the test vessels which are subjected to continuous agitation. The solvent is then removed, if possible completely, by agitation before the test medium is added. The solvent should be non-biodegradable and non-toxic to bacteria, especially when it cannot be removed sufficiently.
- Some sparingly soluble organic compounds dissolve more readily in water when alkali or acid is added. They may be introduced as an acid or alkaline stock solution, provided that no substantial reaction of the test compound takes place. The test medium is adjusted to neutral before the inoculum is added.[14] Based on the Henderson-Hasselbach equation, alkalinisation could be successful for compounds with pKa values for *acidic functionality* larger or equal to 7 (pKa 7 ➔ half of the functional groups will be dissociated at pH 7, ca. 90% will be dissociated at pH 8, increasing polarity and thus solubility), while acidification could be useful for compounds with pKa-values for *alkaline functionality* of smaller or equal to 7 (pKa 7 ➔ half of the functional groups will be dissociated at pH 7, ca. 90% will be dissociated at pH 6, increasing polarity and thus solubility).
- Ultrasonic dispersion at approximately 20 kHz for 30 minutes followed by settlement for 15 to 30 minutes and TOC measurements of the stabilized solution to be used in the test.

---

[12] There is no uniform definition of poor water solubility available: REACH guidance R.7b refers to OECD (< 100 mg/L) and EU-TGD (2003) (< 1 mg/L), outlining that rather below 1 mg/L problems are expected to occur with regard to environmental fate and ecotoxicity testing,

[13] ISO 10634:1995 Water quality -- Guidance for the preparation and treatment of poorly water-soluble organic compounds for the subsequent evaluation of their biodegradability in an aqueous medium

[14] It should be noted that biodegradation tests require a physiological tolerable pH range which could be negatively influenced by the addition of alkali or acids. Furthermore, the test substance may precipitate after addition to test vessel with neutral pH.

- Dispersion with an emulsifying agent, e.g. Synperonic PE/P94, Synperonic PE/P103 or Tween 85, with the limitation that for a valid test degradation extent in the corresponding controls may not exceed 10% of degradation observed in the test item flasks.[15]

The REACH guidance R.7b recommends considering (additional) blank controls treated equal to the test vessels as well as a poorly soluble positive control. Direct addition, particularly via direct weighing or pipetting, or using a support should act as a bench mark for the assessment of all poorly water-soluble compounds. An important draw-back mentioned in the text is the lack of validated and accepted, poorly soluble, but readily biodegradable reference compounds to be included as positive control to check if the methodology is working as desired.

The advantages and disadvantages of the different methods have been assessed by Mead (2014):[16]

The ultrasonic method may be suitable for solids and powders, but seems often not effective for low density oily substances which just float back to the surface. Adsorption to silica gel results in a homogenous dispersion while maximising bioavailability of the test substance. However, low density liquids tend to float off and some chemicals may bind so strongly to the silica that their bioavailability is reduced. Direct addition to an inert support is ideal for highly viscous test items but not for powders which may slide from the support. The use of volatile solvents allows the addition of small amounts of the test substance but is not suitable for volatile compounds and may additionally cause problems with residuals from the solvent. The use of non-volatile and non-degradable solvents may increase the surface area of the test compounds but for many chemicals no suitable solvent is available (Mead 2014)[16].

When testing solid materials with low water solubility it is generally recommended to use appropriate reference compounds. For example, for degradability testing of plastic material according to ISO 14852 microcristalline cellulose powder or polyhydroxybutyrate is recommended as reference material. The ISO standard recommends a maximum particle size of the test and reference species of 250 µm. ASTM D5864, describing a CO_2-evolution test for biodegradability testing of lubricants, recommends rapeseed oil as reference material.[11]

3.1.5.3 Volatile substances

Volatile substances may be removed from open test systems, but also closed systems with a headspace, such as the MITI(I) test (OECD 301C), the Manometric respirometer test (OECD 301 F) or the CO_2 Headspace test (OECD 310). Therefore, these tests are not suitable for highly volatile substances. Volatility is generally described by the Henry air/water partition coefficient H (Atm m^3 mol^{-1} equivalent to 1013 hPa m^3 mol^{-1}). In order to decide whether these methods can be applied the following guidance exists:

- The OECD 310 CO_2-Headspace test (2014, paragraph 11) indicates that test substances up to a Henry coefficient of 50 Pa m^3 mol^{-1} can be tested with that method. This is based on theoretical assumptions, resulting in less than 1% of these substances being expected in the headspace, provided that the recommended headspace to liquid ratio of 1:2 is met.
- When testing potentially volatile substances it has been suggested that a moderately volatile reference item should be used instead of non-volatile ones, in order to assess the influence of the test system. Comber and Holt (2010) recommended using 1-octanol as a relatively poorly

[15] According to Hydrotox's laboratory experience with these methods Tween 85 is not suitable as an emulsifying agent because it is considerably biodegradable itself. ISO 10634 is currently being revised while considering more practical experience with these methods.

[16] Mead, C. (2014). Improving biodegradation of low solubility chemicals: What can we do? Harlan CRS, SETAC Poster Basel 2014.

water-soluble (540 mg/L) reference chemical, which has a moderate volatility and is readily biodegradable in the closed CO_2-Headspace test according to OECD 310. Unpublished results from the Hydrotox laboratory revealed, that 1-octanol is also readily biodegraded in the aerated CO_2 evolution tests according to OECD 301B.

3.1.5.4 Testing of inhibitory substances

Ready biodegradability testing requires relatively high testing concentrations of 2-100 mg/L, which for certain compounds might cause inhibitory effects to the inoculum, resulting in potentially false positive persistency results (i.e. biodegradable substances not recognised as such). If inhibitory effects cannot be excluded OECD 301 (adopted 1992, paragraph 25; OECD, 2016) suggests that a toxicity test, containing both the test substance and a reference compound, should be considered. If in this inhibition control *"less than 35% degradation (based on total DOC) or less than 25% (based on total ThOD or $ThCO_2$) occurred within 14 days, the test substance can be assumed to be inhibitory."* In this case *"the test should be repeated, using a lower concentration of test substance ... and/or a higher concentration of inoculum, but not greater than 30 mg solids/l"*.

These criteria are based on the assumption, that the reference compound contributes to about 50% of the total DOC, ThO_2 or $ThCO_2$ introduced to the inhibition controls and is biodegraded for less than 50% within 14 days. This relatively rough estimate may be supplemented by further criteria such as negative degradation extents observed in the inhibition controls at the beginning of the tests, which is an indicator that the inoculum was less active and therefore inhibited at this stage.

Wess et al. (2014)[8] stated that the OECD criteria for inhibitory effects may not indicate the level for relevant inhibition and suggest that when a repetition of the test with lower concentration shows higher biodegradation, this result should overwrite a former result of a partly intoxicated inoculum.

ANNEX II of OECD 301 refers to options for testing ready biodegradability of chemicals suspected to be toxic to the inoculum. In order to avoid inhibition due to toxicity it is suggested that the test substance concentrations "should be less than 1/10 of the EC_{50} values (or less than EC20 values) obtained in toxicity testing. ... EC50 values of less than 20 mg/l are likely to pose serious problems for the subsequent testing. Low test concentrations should be employed, necessitating the use of the stringent and sensitive Closed Bottle test or the use of 14C -labelled material. Alternatively, an inoculum previously exposed to the test substance may permit higher test substance concentrations to be used. In the latter case, however, the specific criterion of the ready biodegradability test is lost."

In principle, the Closed Bottle test (OECD301 D) has the lowest test concentration (about 2 mg/L) of all OECD 301 tests and thus might be suitable for inhibitory substances. Unfortunately, the method has also the lowest biodegradation potential (Figure 1). Thus, in the Hydrotox laboratory often the lower concentration of the CO_2 evolution test (OECD 301 B) of 10 mg TOC/L is used for testing potentially inhibitory substances. A comparison of the test results obtained at 10 mg TOC/L with that obtained from a parallel series with the standard concentration (20 mg TOC/L) gives an indication of inhibitory effects and their influence on biodegradation. Test concentrations below 10 mg TOC/L are not recommended when using the standard inoculum concentration of 30 mg d.s. /L activated sludge, because the difference to the inoculum blank is not sufficiently large and might become insignificant.

3.2 Screening tests for inherent biodegradability

According to OECD (OECD/OCDE, 2006) the tests of inherent biodegradability are designed to assess whether the chemical has any potential for biodegradation under aerobic conditions. Inherent biodegradability is measured by specific analysis (primary biodegradation) or by non-specific analysis

(oxygen consumption, DOC elimination).[17] According to OECD biodegradation extents above 20% of theoretical values (measured as BOD, DOC removal or COD) may be regarded as evidence of inherent, primary biodegradability, whereas biodegradation above 70% of theoretical (measured as BOD, DOC removal or COD) may be regarded as evidence of inherent, ultimate biodegradability.

The Zahn-Wellens test according to OECD 302B is the inherent test most often applied. The test uses relatively high test concentrations (50-400 mg/L DOC) as well as high inoculum concentrations (200 – 1000 mg d.s./L activated sludge). Measured degradation is compared to the values measured after 3 hours (reference point), assuming that all adsorption processes are completed after this time. When the DOC difference between the test vessels measured after 3 hours is unexpectedly low (e.g. in the same range as in the inoculum blank vessel), this is an indicator for physico-chemical adsorption. In such cases the DOC or COD after 3-hours-should be compared with the values measured before the inoculum is added (OECD 302B point 29). This amount is reported as "adsorbed by the activated sludge". In routine testing this information may be lost, because the values measured after 3 hours are taken as start concentration.

ISO 9888 provides further guidance on the interpretation of the Zahn-Wellens test results: If the DOC- or COD-concentration after 3 hours is significantly lower than at the start (> 20% elimination), this is interpreted as abiotic elimination due to adsorption or volatilization. In this case ISO 9888 states that the total elimination of the test substance should be indicated as additional information. For calculating the total elimination the (calculated or measured) start concentration is taken as reference point instead of the 3-hour value.[18]

To be used in the persistence assessment according to REACH, REACH guidance document R.11 (ECHA, 2014c) is even stricter in this respect, insofar that the removed fraction before degradation occurs (usually 3-h value) must remain below 15% in a Zahn-Wellens test according to OECD 302B (for details, see section 4.2.2).

CONCAWE (Battersby et al., 1999; CONCAWE, 1999) developed an inherent biodegradation standard test protocol especially for poorly soluble, volatile oil products. Based on International Standard IS0 14593 (CO_2, Headspace Test) it differs mainly in the use of a 14 day pre-exposed (to the test item, i.e. adapted) inoculum and extended test duration up to two months. For this, inoculum samples were taken from sites previously exposed to oil products (refinery biotreater, contaminated soil). The inoculum was used at a final concentration of 10%. The test items were dosed by direct weighting via a glass fibre filter solid support. Results of an international ring test with 12 participating laboratories and four oily or waxy y test items and the reference item Hexadecane were reported. The evaluation resulted in relative high coefficients of variance (e.g. a CV of 21% was determined for the reference item).

Beek et al. (2000) stated[19] that the Zahn-Wellens test (OECD 302B) and the SCAS test (OECD 302A) do not distinguish between biological degradation and other elimination mechanisms, because the DOC analytic does not discriminate between both. They suggested that this should be taken into account by the supplement "biodegradable/eliminable". Currently, the ECHA guidance refers to both tests with the term "degradability". Further on, the SCAS test is not considered comparable with

---

[17] The OECD guidance attributes the term "ultimate biodegradation" to all these non-specific analyses, which should exclusively be used to methods determining mineralization (oxygen uptake or CO_2-evolution).

[18] ISO 9888:1999 Water quality -- Evaluation of ultimate aerobic biodegradability of organic com-pounds in aqueous medium -- Static test (Zahn-Wellens method)

[19] Chapter *The Assessment of Biodegradation and Persistence* (Beek, B., Böhling, S., Franke, C., Jöhncke, U., Studinger, G., Thumm, E.)

other tests for inherent biodegradability due to the test design, which considers a discontinuous operation modus, a high inoculum concentration, nutrient addition, and a long adaptation phase. This is consistent with the REACH guidance R.7b which does not accept data from the SCAS test for determining inherent biodegradability.

Several approaches have been performed to combine DOC-elimination with an additional endpoint for ultimate biodegradation (Baumann and Müller, 1996; Gartiser et al., 2007; Jiang et al., 2002; Meinecke et al., 2000; Strotmann et al., 1995). A screening test combining DOC removal and CO_2-evluation measurements, optimized with regard to inoculum concentration and composition of medium, was elaborated by Strotmann et al. (1995). A comparison of results with outcomes from OECD 301F and OECD 302B led the authors to judge this test as giving reliable results and providing enhanced information on the pattern of biodegradation. This approach has been further developed for difficult substances by Gartiser et al. (2007). Here, the modified Zahn-Wellens test (OECD 302B) was combined with devices to monitor mineralization (CO_2-production), as it is realised in the CO_2 evolution tests according to OECD 301B. Thus, in addition to the DOC elimination also CO_2 evolution is measured in order to determine inherent biodegradability. Therefore, with this test, adsorption (by decline of DOC) and mineralization can be monitored simultaneously, such that a conclusion may be drawn if and to what extent degradation in the sorbed state is possible. Although not being standardised so far, this test may serve as equivalent to the MITI (II) test according to OECD 302C). However, while oxygen consumption is an indirect indicator for the mineralization of a substance and might be influenced also by the oxygen supply for nitrification, CO_2 evolution constitutes an unambiguous evidence of ultimate biodegradation (Gartiser et al., 2007).

The differentiation between the adsorbed fraction and mineralisation also has regulatory consequences, because for compounds with a high adsorption potential, often low or no degradation is assumed by default for the adsorbed fraction. Due to the relatively high sludge content, the elimination after three hours till 24 hours in the OECD 302B is accepted as a screening method for adsorption according to REACH guidance document R.7a (ECHA, 2014a).

3.3 Enhanced screening tests

3.3.1 Overview

Enhanced screening tests are essentially derivatives of the OECD 301 test series with introduced modifications facilitating biodegradation. They are meant to "improve the environmental relevance of biodegradability assessments without the immediate requirement for simulation level testing" (ECHA, 2014b) and may be used to demonstrate non-persistence with regard to PBT/vPvB assessment (see section 4.2.2). In REACH guidance R.7b (ECHA, 2014b), the following possible modifications are suggested:

- Prolongation of test duration up to 60 days, especially for poorly soluble substances.
- Use of larger test vessels to increase microorganism diversity and absolute numbers without changing inoculum density: higher probability for presence of competent microorganisms.
- Increasing the biomass concentration and/or testing at different biomass concentrations using concentrated microbes of environmental waters. This approach may allow conclusions on the volume of environmental waters (e.g. river water) which would be needed to provide sufficient microorganism diversity to enable test item degradation (following the most probable number approach).
- Using low level pre-adaption: because adaption and enrichment phenomena are naturally occurring in the environment, a suggestion is to use the inoculum of a first ready biodegradability test in a subsequent further ready biodegradability test.

- Using semi-continuous adaption/selection of inoculum: inoculum for a ready biodegradability test is derived from a test system fed with the test item at environmentally realistic (low) concentrations in a semi-continuous manner, allowing adaption/selection over time whilst leaving the diversity, viability and nutrient status of the test system largely undisturbed. An example for such a system is the semi-continuous version of OECD 309.

These modifications mentioned in R.7b are meant as suggestions; thus, further modifications may be acceptable if the prerequisite of these tests – environmental realism and relevance – is maintained.

3.3.2 Inoculum

3.3.2.1 Inoculum source

It is recognized that the inoculum used in standard laboratory screening tests only represents a small range of environmental complexity and heterogeneity. For example, microbial biofilms are not used in screening studies. Ready biodegradation tests mainly detect growth-linked biodegradation by using the test item as only substrate at a high concentration. For some test items such as biocides the level of concentration might however inhibit the inoculum (ECETOC, 2013c).

Inoculum biomass and diversity is known to be the greatest source of variability of test results. Inoculum density in ready biodegradation tests (from the Closed Bottle test (OECD 301 D) to the CO_2-Evolution test (OECD 301 B) varies by several orders of magnitude (10^4 to 10^8 cells /L) and is far from being standardised (ECETOC, 2003). The likelihood of competent degraders being available depends on the inoculum source and the inoculum density. The last is mainly determined by analytical and testing conditions as well as the discriminative power of the measurements in test and control vessels. In the Closed Bottle (OECD 301 D) test the oxygen solubility in water determines the inoculum density. Other tests (e.g. CO_2 headspace test (OECD 310) or Closed Bottle test (OECD 301D) require a reduction of the organic carbon introduced with the inoculum through washing in mineral medium or acclimation (ageing) of the inoculum in order to fulfil the validity criteria of the respective guidelines. However, no consistency exists among different testing schemes. In this respect, an intensive review was published recently (Kowalczyk et al., 2015), critically investigating constraints of current screening tests (ready biodegradability test, enhanced screening tests, inherent biodegradability tests). An update of current testing guidelines is suggested, mainly with regard to a better characterisation and standardisation of inocula (see also sections below). To this end, beyond others the application of high throughput combined with methods to characterize and quantify pools of biological molecules ("OMICS" methods) are suggested. From this, better insight into biodegradation pathways of chemicals as well as their environmental fate is anticipated by the authors.

The variability of the inoculum in terms of quality and quantity has been described to mask all other factors in biodegradability testing especially at low biomass concentrations. This has been assessed by Blok et al. (1984) via a theoretical simulation of the Monod growth kinetics, depending on the number of competent bacteria at the start, corrected for cell decay. The basic assumption was that the number of bacteria able to degrade a specific chemical as only carbon source will be high for chemicals such as glucose, but may be very low within the same inoculum for chemicals with low biodegradation potential, such as tertiary butanol. The expected degradation extent (as oxygen demand) corresponds to the elimination through metabolism less the cell growth due to the degradation of the chemical, which depends on the specific growth rates. From these simulations the authors concluded that variability of results between different test runs and laboratories can be completely explained by variation of the inoculum. These assumptions were also confirmed by experimental results while using different substrate/inoculum ratios (ECETOC, 2003; Vazquez-Rodriguez et al., 2000).

On the other hand, some authors assume that activated sludge from STPs of similar configurations and designed for treating wastewater of similar compositions will roughly have a similar microbial community structure (Seviour et al. 2010, cited after Vázquez-Rodríguez et al., 2011).

There have been several attempts to standardise the inoculum source in order to enhance comparability of test results. Paixão et al. (2006) compared the degradation behaviour of the reference compound diethylene glycol in the Zahn-Wellens test (OECD 302 B) with two different activated sludges and three defined microbial consortia (two commercially available and one composed of six selected bacterial strains commonly found in sewage). While the two activated sludge inocula resulted in the best degradation extents and revealed a good comparability and repeatability of results, the three standardised inocula only reached the pass level of 70% COD-elimination within 14 days when sterilised supernatant of activated sludge was added to the test mixtures. Although the experimental data show that the attempt to standardize the inoculum resulted in lower biodegradation extents the authors concluded that designed inoculum may be an alternative to activated sludge.

It has further been reported that also the mass loading rates (MLR) of treatment plants (g BOD_5 per g dry solids and day) has a decisive influence on the biodegradation potential. The sludge with the highest mass load had the highest activity in terms of respiratory activity, cultivable cells and the hydrolytic enzyme profile (Vázquez-Rodríguez et al., 2007). Vázquez-Rodríguez et al. (2003; 2011) studied the degradation behaviour of several chemical substances[20] with the OECD 301 A test inoculated with the supernatants from activated sludge collected in three STP operating at low, mean, and high MLR (0.1, 0.5, and 0.9 g BOD per g dry solids and day, corresponding to sludge retention time (SRT) of 30, 5, and 3 days respectively). The higher the MLR, the higher was the activity of the inoculum in terms of colony forming units and the shorter were the lag phase before the degradation started. The inoculum activity of the supernatants (BOD_5 from 25-125 mg/L) correlated with the suspended solids (20-155 mg d.s./L) and the cultivable active cells (10^7-10^9/L). The authors concluded that activated sludge from STP operating at high MLR has a higher biodegradation potential than that operating at the lower MLR.[21]

Similarly, Struijs et al. (1995) analysed the oxygen consumption of activated sludge and secondary effluent from 40 STP in the two phase closed bottle test and closed bottle test, respectively and found a strong dependence on the mass loading rate for both inoculum sources. This study was not complemented with real degradability studies.

On the other hand, there are also examples where lower MLR improved the biodegradation behaviour. For example, the removal rate of the X-ray contrast agent Iopromide and the antibacterial drug Trimethoprim was significantly higher in nitrifying activated sludge, compared to activated sludge whose nitrifying bacteria were inhibited (Batt et al., 2006). Similarly, Torres-Bojorges and Buitrón (2012) tested the primary biodegradability of a technical mixture of nonylphenols (tNP) with three different inocula in batch tests. Nitrifying sludge presented the highest biodegradation percentage compared to fresh activated sludge or sludge pre-exposed to 4-chlorophenol.

Because the MLR also determines the sludge retention time (SRT, also called sludge age)[22] it also has an influence on the microorganism groups which may establish in the activated sludge. For example,

---

[20] Dodecyl benzene sulfonate, Nitrilotriacetate acid (NTA), Pentaerythritol, Sodium acetate, Aniline.

[21] It should be noted that inoculum from STP with high MLR and a higher microbiologic activity also will result in higher inoculum blank values.

[22] Shortly, the sludge age (or sludge retention time) is the ratio of the activated sludge (kg dry solids) in the aeration basin to the excess sludge removed (kg dry solid per day). The sludge age describes the mean retention time of activated sludge in STP.

the nitrifying bacteria *Nitrobacter* and *Nitrosomonas* do not grow sufficiently in activated sludge below a sludge age of approximately 6 days, because of their lower growth rate (Wiesmann et al., 2007).

Summarising, the MLR of the inoculum source has a decisive influence on the biodegradation potency and on the inoculum blank. When using inoculum derived from activated sludge with a low MLR (or high SRT) the reduced activity in the inoculum blank may be compensated by higher inoculum concentrations.

For a better harmonisation of the activated sludge inoculum it could be prescribed to give an indication of the MLR in order to assess whether the sludge has some nitrifying capacity. Normally, the European urban waste-water treatment Directive 91/271/EEC demands a nitrogen reduction of 75-80% for STP with more than 10.000 population equivalents[23] and thus the size of the STP does already provide an indication of the nitrifying capacity.

In general, it is assumed that activated sludge is a very competent inoculum due to the high cell concentration. This inoculum is often pre-treated by washing, filtration or centrifugation in order to reduce the DOC introduced into the tests. Hereby, bacteria floating in the water phase and not attached to the activated sludge are removed to a certain extent, depending on the pre-treatment (see chapter 3.3.2.2).

Most experiments cited above used activated sludge supernatant (partly previously sonicated), but there is little information whether settled activated sludge (dry solids), which is the standard inoculum in many OECD 301 test, has a similar dependence on the MLR and the biodegradation potential like the supernatant inoculum. In fact, both fractions may have their pros and cons as degradation agents in testing due to their different behavior in STP: while the residence time for water soluble chemicals and water-borne bacteria is mainly determined by the hydraulic residence time in STPs (ca. 6 h) that of insoluble chemicals adsorbed to activated sludge and the bacteria therein is mainly determined by the sludge retention time (sludge age ca. 5-10 d) (e.g. Cowan et al., 1993; Wiesmann et al., 2007). Thus, supernatant and settled activated sludge are expected to contain bacteria with different growth kinetics.

Usually the supernatant of activated sludge, which in fact corresponds to secondary effluent, is disregarded through these washings. The OECD 301 A allows up to 10 % secondary effluent as inoculum. This amount corresponds to the situation for surface water in central Europe, where secondary effluent represents about 10% of river water flow according to ECHA guidance documents. Obviously the different pre-treatment procedures for the activated sludge cause different efficiencies for the removal of the supernatant. (Settlement of activated sludge followed by resuspension in dilution water is less effective than centrifugation or filtration of activated sludge followed by resuspension.)

One option for a better harmonisation of the inoculum source and improving the inoculum potency without counteracting the requirements for inoculum control pass levels could be to separate the solids and supernatant of activated sludge followed by a defined reunion of both sources (e.g. allowing 10% of supernatant contained in the re-suspended activated sludge).

3.3.2.2 Pre-treatment and characterisation of the inoculum

The inoculum concentration in ready type tests often is limited in order to avoid overly high (and scattered) background values in the inoculum blank vessels, which are subtracted from measured

[23] The population equivalent of one person corresponds to the organic biodegradable load having a BOD_5 of 60 g of oxygen per day.

test values and could reduce the discriminative power of the test. As a consequence several approaches have been undertaken to reduce the background level of organic carbon in inocula in order to reduce the blank values and to obtain a better characterised inoculum. In principle, these pre-treatment steps also aim at obtaining a standardized inoculum.

Washing with mineral medium, concentration (centrifugation, filtration), colonisation on glass beads and/or pre-incubation of the inoculum for 5-7 days has been used for reducing the blank values. It can be assumed that the major part of total blank respiration during the first week is due to the oxidation of adsorbed organic ingredients (Struijs et al., 1995). However, most preconditioned inocula have reduced activity and potency. Some authors concluded that specialist populations (e.g. aniline degraders) are more sensitive to pre-treatment steps than generalist populations (Vazquez-Rodriguez et al., 2000; Vázquez-Rodríguez et al., 2007).

The effect of inoculum pre-treatment through washing and a 7 day pre-incubation was tested by Vázquez-Rodríguez et al. (2007), who used as inoculum the supernatants of different activated sludge sources obtained from slight centrifugation (4 minutes at 500 g). The preconditioning led to a diminution of the biodegradation potential for the readily biodegradable compounds sodium acetate and aniline. The reference substance aniline was not being degraded anymore within the observation period of 90 hours. The authors concluded that methods intended for biomass homogenization must be further developed.

Thouand et al. (2011) partly built on results of Vázquez-Rodríguez et al. (1999), elaborating that one of the most important parameters determining results of biodegradation screening tests is the ratio of substrate to biomass. This ratio was found to be determining lag time, biodegradation levels as well as final carbon distribution in cellular, mineralized and residual carbon. The authors recommend either a quantitative determination of this ratio or fixing it at a certain value for standardized ready biodegradation tests to enhance reliability of test results. The importance of this ratio is further corroborated by more recent work in this regard (Vázquez-Rodríguez et al., 2006).

Goodhead et al. (2014), using a high-throughput screening test in 96-well plates, found that the treatment of activated sludge inocula through sedimentation or filtration drastically reduced the number of bacteria and the overall diversity of bacteria compared to the original samples. These effects are detrimental to bacterial community structure and reduce cell numbers as well as operational taxonomic unit richness. The authors could further demonstrate that this was associated with a significant reduction of degradation probability and a corresponding increase of variability of biodegradation of 4-nitrophenol, compared to the use of unprocessed inocula. When inoculated with the standard amount described by OECD 301 of 30 mg/L dry solids or 1 ml /L effluent, no biodegradation of 4-nitrophenol at all was detected. Unprocessed activated sludge only at much higher inoculum concentrations was able to biodegrade 4-nitrophenol, while degradation by settled or filtered inoculum failed. It should be noted, that these results may also be explained with the low test volume in the 96 well plates compared to standard OECD 301 tests, which also reduces the absolute number of competent bacteria at the start. The authors themselves confirm that the absolute number of bacteria in typical OECD tests would exceed those used in their study, if they did not increase the inoculum concentrations.

Existing biodegradability data were evaluated recently by comparing OECD 301C (MITI I test) with other OECD 301 tests (Kayashima et al., 2014). The main difference of OECD 301C is inoculum pre-incubation with synthetic sewage (containing glucose and peptone), while all other tests are using activated sludge or secondary effluents as inoculum source. It turned out that biodegradation potency of OECD 301C is weak compared with tests using the sludge directly.

In a review on standardization of activated sludge for biodegradation tests, Vázquez-Rodríguez et al. (2011) comprehensively summarize factors influencing microbial composition and diversity of

sludges which may have an impact on biodegradation potential. This includes microbiology of sludges, selective pressures, different sources, pre-conditioning, pre-culturing, and preceding adaptation of sludges. The authors conclude that "any manipulation of the inocula incurring in a diminution of their biodegradation potential, for example preconditioning or pre-culture, does not constitute an option for standardization", as an increase in reproducibility is achieved at the cost of microbial diversity, the latter being essential to reliably assess environmental biodegradation potential of compounds.

Other approaches consider homogenisation of the inoculum. Thouand et al. (1996) sonicated river water inoculum. Foladori et al. (2007) concluded that while bacteria in general are quite sensitive to sonification, activated sludge samples disaggregated at low sonification levels, releasing single cells in the bulk liquid, while disruption of bacteria was induced only by very high sonification levels. This means that any pre-treatment through sonification should be carefully performed and the effect on viability should be analysed beforehand.

It has been suggested that pre-incubation of the inoculum could also reduce its background activity with the aim to allow higher inoculum densities, because the favourable signal to noise relation of lower densities can be maintained also for higher densities by such an approach (ECETOC, 2007). This is further discussed in chapter 3.3.2.3.

There have often been complaints that the inoculum variability allowed in the OECD tests is responsible for conflicting results of different tests (Kowalczyk et al., 2015). The authors concluded, that standard ready biodegradability OECD tests are not fit for prioritization of chemicals based on persistence (Goodhead et al., 2014). On the other hand, the OECD 301 states that the inoculum source and test conditions should be carefully analysed when conflicting results have to be discussed and available data from literature do not support the selection of a specific standardized inoculum due to the drawbacks described above. As a consequence, there have been several attempts to better describe the microbiological characterisation of the inoculum at the start and during the test. One step in this direction is the qualitative characterization and estimation of the relative composition of bacterial communities present in normal inocula as well as specifically considering certain assemblages responsible for certain critical degradation pathways.

In this respect, microbial diversity of a mixed culture capable of de-chlorination of 1,2-dichloropropane was analysed by Schlötelburg (2001). Several molecular genetic methods were used to accomplish this. As a future perspective knowledge on specific assemblages of species responsible for essential biodegradation pathways combined with knowledge on what may usually be present in environmental media or STP sludges could lead to preserved mixed inocula from the shelf, most probably enhancing reproducibility and representativeness of biodegradation tests.

One important point is that the inoculum source and the effect of pre-incubation (or pre-exposure to the test item) should be described with accompanying analysis. A relatively simple routine analysis at treatment works is light microscopic analysis of the activated sludge for characterising the inoculum. Other methods applied with the aim of describing the bacterial diversity or changes of the community structure resulting from any pre-treatment of the inoculum or during the test duration are e.g. cell counting combined with Denaturating Gradient Gel Electrophoresis (DGGE) of DNA (number and pattern of bands identified as indication for diversity) (Martin, 2014).

The most ambitious, but also most costly approach for describing the influence of the diversity of the inoculum on degradability would be to identify different phylogenetic groups and strains by in situ hybridisation using specific oligonucleotides (16S rDNA) for determining species richness and distribution (Snaidr et al., 1997; Wenzel, 2002). This would allow the calculation of biodiversity indices as is routinely used for ecology analysis. However a complete analysis costs several thousands of EUR. As an alternative the main bacteria groups can be analysed with a set of phylum-specific probes in

order to obtain a general overview of the bacterial groups present. Similar approaches are described by other authors (Bartram et al., 2011; Goodhead et al., 2014). Well established next-generation sequencing techniques facilitate this detailed analysis and even allow for monitoring potential changes in microbial diversity at different time points.

While providing no information on inoculum diversity, several conventional methods for characterization of inoculum activity have been described by Wos (2005). These include ATP and NADH measurements (evidence of energy status of cells), uptake of thymidine into DNA (evidence for cell synthesis), uptake of leucine into proteins (evidence of protein synthesis), oxygen uptake (evidence of respiration), or DOC measurements (evidence of substrate assimilation). Several authors used direct cell count of bacteria (by epifluorescence) and determination of cultivable bacteria (e.g. Thouand et al., 1995). It is known that only a small part of the bacteria present in activated sludge can be cultivated with nutrient agar or other substrates such as peptone, yeast extract and others. Thus the use of a special activated sludge agar (using activated sludge as part of the substrate) has been suggested by Kappesser and Kutzner (1991). Other endpoints for describing the potency of the inoculum used include the determination of the dehydrogenasic activity or the profile of hydrolytic enzymes (Vázquez-Rodríguez et al., 2007).

Additionally, the behaviour of test systems exposed to reference compounds (lag phase, degradation extent) might provide information on the potency and versatility of an inoculum. For this some reference compounds with slower degradability than sodium acetate or sodium benzoate (e.g. diethyleneglycol or other substances which are degraded in the Zahn-Wellens test but often fail the pass criteria of ready biodegradation tests) could be used. From literature several chemicals with these characteristics have been described (see chapter 3.3.3).

Summarising, the main objective of the pre-treatment of the inoculum is to fulfil the requirements of the test guidelines with respect to the inoculum blank controls. Available literature data show that any pre-treatment of the inoculum by pre-incubation, filtration, centrifugation, or sonification with the aim of standardising the inoculum source reduces the activity and potency of the inoculum. The inoculum source, its activity and diversity could better be described by means of microbiological and molecular techniques, in order to understand the biological background of biodegradation.

3.3.2.3 Increase of inoculum biomass

As discussed in chapter 3.3.2.2 the biodegradation success in a specific test is mainly dependent on the number of competent bacteria present at the start of the test. This number could either be increased by elevating the inoculum concentration or by increasing the flask volume of the test. While the inoculum concentration has its limits set in the OECD 301 guideline for ready type degradation tests, an increase of the flask volume should in principle be acceptable without questioning the attribution as "ready type" biodegradation test. For example Mead et al. (2013)[24] used an enhanced OECD 301 B test with 4000 mL liquid volume while Menon (2014) used a Closed bottle test with 1000 ml volume.

When it is intended to increase the inoculum concentration in order to perform an enhanced screening biodegradability test it has to be decided which inoculum concentration might be acceptable. There is a gap between ready biodegradation and inherent biodegradation type tests (inoculum density up to 30 mg/L dry solids for ready type tests and 200 mg/L as lower level for inherent tests).

---

[24] Mead, C. Clarke, N., Bayliss, B. (2013). Enhanced biodegradation tests; Application to persistency evaluations. Harlan (Poster) http://www.harlan.com/download.axd/640e145dd21b429aa2d5df6df9c65fd9.pdf?d=Enhanced Biodegradation Tests_Application to Persistency Evaluations

Test guidelines such as ISO 14852 on biodegradability testing of plastic material based on the CO_2 Evolution test allow a wide range of test concentrations (100-2000 mg C/L) and inoculum densities (30-1000 mg/L) in order to address the low water solubility and bioavailability of the test item.[25] This could easily be adopted also for chemicals (other than polymers) with low water solubility. The higher inoculum blank values expected in these tests are balanced with higher testing concentration, which for chemicals with low water solubility normally is no problem due to their lower bioavailability and therefore limited toxicity to the inoculum.

The effect of dilution on microbial communities has been studied by Franklin et al. (2001) by numerical simulations and batch culture experiments. In theory, a dilution of a relatively diverse community would remove rare organism types, creating mixtures of cells differing in species richness. The results of the numerical simulations showed that while microbiologists generally consider dilution to be a linear process, the response of various community level parameters (richness, evenness, and diversity) to such a manipulation may produce nonlinear results leading to a rapid loss of species richness depending on the variance of the distribution of individuals. In theory, also biomass growth of a community with high diversity (and interspecific competition) is lower compared to communities with lower diversity and less interspecific competition. These aspects have been proven in batch experiments. This could explain, why the plateau phase of reference substances tested in the Closed Bottle test (low inoculum density) according to the experience of the Hydrotox laboratory reaches about 70-80%, while in the CO_2-evolution or Respirometer tests (with relatively high inoculum density) usually values above 90% are observed. The reason is that the carbon used for biomass growth is fixed in cells and thus cannot be detected as oxygen consumption or CO_2 evolution. The extent of biodegradation of a suitable reference compound, which to date only is used as a validity criterion, could thus also provide further useful information about the potency of the inoculum.

Some important questions with regard to possible enhancements of existing biodegradation tests are addressed in a recent PhD-work (Martin, 2014). In this study the applicability of enhanced screening tests with positive and negative reference compounds as suggested by Comber and Holt (2010) was assessed with an OECD 301B type test at different activated sludge concentrations (0.3 – 3000 mg d.s./L) while measuring the evolved $^{14}CO_2$. The test chemicals used were expected to pass ready type biodegradation tests (aniline) or to be only biodegraded under enhanced conditions (4-nitrophenol, 4-fluorophenol, 4-chloroaniline), Further on, a negative control substance (pentachlorophenol) not expected to be biodegradable - even under enhanced conditions - was tested. With regard to activated sludge based inocula, higher cell densities resulted in higher degradation rates and shorter lag-periods, but the most pronounced effect was a decrease in inter-replicate variation. The effect of inoculum concentration was higher than that of test flask volumes. All enhanced biodegradation tests sufficiently degraded 4-nitrophenol to 60% within 28 days and a 10-day window. The degradation of 4-fluorophenol reached 50-63% across the range of inoculum concentrations applied. As expected, even under enhanced test conditions, pentachlorophenol was not removed. Thus, no false positives for biodegradability (i.e. non-biodegradable substances identified as biodegradable) were observed within these enhanced test systems. Microbial diversity was observed to increase with increasing cell densities. Concomitantly, similarity with regard to the microbial assemblies increased between replicates. In contrast, upon increase of cell density for marine systems, no such effects could be observed.

The following options for allowing higher inoculum concentrations without counteracting the general principles of the test categories may be discussed:

---

[25] ISO 14852:1999 Determination of the ultimate aerobic biodegradability of plastic materials in an aqueous medium -- Method by analysis of evolved carbon dioxide.

- Allow higher inoculum concentrations as far as the validity criteria for the inoculum blanks are fulfilled. For example, when the inoculum activity is reduced by pre-incubation to the test conditions, activated sludge concentrations above30 mg d./L may be allowed e.g. in the OECD 301 B CO2 evolution test as far as the inoculum blanks are still below 40 mg/L respective 70 mg/L of CO_2.
- Allow higher inoculum concentrations as far as the test category of ready or inherent tests is still maintained. For example higher inoculum concentrations than 30 mg d.s./L activated sludge may be acceptable for the ready test category as far as the distance to the lower inoculum concentration for inherent test of 200 mg d.s./L is kept (e.g. 50% of 200 mg d.s. /L, corresponding to 100 mg d.s./L).

3.3.2.4 Adaption to test item

The DOC based ISO 7827[9] notes that under certain conditions adapted (ISO uses the term pre-exposed) inoculum may be used, provided that this is clearly stated in the test results (e.g. % biodegradation, using pre-exposed inoculum).[9] Pre-exposed inoculum can be obtained from laboratory biodegradation tests (e.g. Zahn-Wellens-test OECD 302B and ISO 9888[18] or the SCAS test OECD 302A and ISO 9887 [26]) or from samples collected from locations where relevant environmental conditions exist (e.g. treatment plants dealing with similar compounds or contaminated areas). Similar provisions are also given in other ISO biodegradation standards such as ISO 10708.[27] While the application of the SCAS test for inherent biodegradability testing is not accepted according to REACH (see chapter 3.2) it might be used for obtaining pre-exposed (adapted) inoculum. If pre-exposed inoculum is used, the results are interpreted as "inherently biodegradable" according to ISO.

For test substances with inhibitory effects to the inoculum Annex II of OECD 301 (1992) allows that the inoculum may be pre-exposed to the test substance in order to permit higher test substance concentrations after adaptation. REACH guidance R.7b also refers to the possibility of an initial low–level pre-adaption test followed by a second ready biodegradability test using the inoculum derived from the initial test. Therefore, in case of test items toxic to the inoculum at concentrations necessary for biodegradation screening tests, a pre-exposure stage at lower concentration might be an option to achieve a decrease of sensitivity of the inoculum and thus to allow for higher test item concentrations required for screening tests.

In principle there might be several strategies for obtaining and using pre-exposed inoculum. Pre-exposure of the inoculum at low concentrations (µg/L) followed by increasing the concentration of the test item/substrate (semi-batch principle within one test), pre-exposure at low concentrations followed by a subsequent ready type test, or collecting adapted inoculum from contaminated sites.

No guidance for performing enhanced biodegradation tests with pre-exposed (adapted) inoculum exists and some proposals might improve their acceptability by regulatory authorities. For example, it has been suggested to compensate the use of adapted (pre-exposed) inoculum by changing the pass level to e.g. > 70% $ThCO_2$ or ThOD instead of the standard pass level > 60% (Bowmer and Leopold, 2004). In fact, own experience at Hydrotox from hundreds of tests with the OECD 301 B and F tests show that degradation for the reference compounds sodium acetate and sodium benzoate is rather in the range of 90% than in the required range of > 60%. This implies that the carbon used for biomass growth is not as important as suggested. As a consequence, an increase of the general pass level for

---

[26] ISO 9887:1992. Water quality -- Evaluation of the aerobic biodegradability of organic compounds in an aqueous medium -- Semi-continuous activated sludge method (SCAS)

[27] ISO 10708:1997 Water quality -- Evaluation in an aqueous medium of the ultimate aerobic biodegradability of organic compounds -- Determination of biochemical oxygen demand in a two-phase closed bottle test.

ready biodegradability to only 70% ThOD or $ThCO_2$ when pre-exposed inoculum is used may not be really protective when being used for P-assessment.

Adaptation techniques to induce/select microbial degradation potential with regard to a specific substance are described by Watson (1993). The so-called "single-flask procedure", where microorganisms are adapted for the test item over 2 to 7 d in a single flask, either at constant or with increasing concentrations, turned out to be more effective than a procedure with successive selection and transfer steps.

Mezzanotte et al. (2005) analysed the influence of different activated sludge inocula on the biodegradability of polycaprolactone and a starch-based material according to ISO 14851[10]. One batch of activated sludge was used before and after a period of acclimatisation to both compounds while being additionally fed with a starch suspension. The adaptation process improved the biodegradation of both materials compared to the original activated sludge, but the sludge from the acclimatisation procedure still did not reach the potency of another sludge derived from industrial bioreactor for treating chemical-pharmaceutical wastewater.

Adaptation to antimicrobials: Adaption potential of microorganisms to degrade a quaternary ammonium surfactant, dodecyltrimethylammonium chloride, was investigated in a model stream dosed with the compound: the ability to degrade the test item increased by 10 to 1000 fold at all downstream locations from the dose site. Further, a resistance to loads normally inhibiting biodegradation activity was observed after prolonged exposure. While a loss of adaptation was observed concomitant to a cease of exposure, a rapid re-adaption was observed upon re-exposure (Shimp et al., 1989).

Adaptation to new chemicals introduced to the market: One example was described by Sparham et al. (2008) who analysed the biodegradation of highly ethoxylated (>20 EO) alcohol ethoxylate surfactants (AEs) and found that the observed ready biodegradability contrasts with earlier biodegradation studies on similar chemicals. This was explained by the authors with a significant adaptation of microbial communities because of the increased discharge of AEs to municipal sewers. However, a test design that purposefully employs adaptation is not relevant for regulatory use, since it would stand in contradiction to the required precautionary character of the assessment.

Importance of the inoculum source/pre-treatment: Strotmann et al (1993) studied the biodegradation of morpholine in several static biodegradation tests and in a laboratory-scale STP. While pre-exposure of the inoculum had no significant influence in the OECD Screening test (301 E) the lag phase was significantly reduced in the Zahn-Wellens test and the laboratory STP. The lag period seemed to be due to the cell number of morpholine degrading bacteria in the activated sludge, which is much higher in a per-exposed sludge than in non-pre-exposed sludge.

These results illustrate the importance of per-exposure in determining the fate of synthetic chemicals in aquatic environments.

Mechanism of adaptation due to pre-exposure: The adaptation process depends on the concentration of the chemical. Spain and Van Veld (Spain and Van Veld, 1983) found a threshold concentration of 10 ppb (µg/L) p-nitrophenol below which no adaptation was detected. The biodegradation rates with pre-exposed inoculum also increased with concentration.

The existence of some kind of "adaption threshold" is corroborated by Toräng et al. (2003), who observed "shifts in biodegradation kinetics" for the herbicides MCPP and 2,4-D at low concentrations in aerobic aquifer materials. At concentrations above 1 µg/L microbial biodegradation was induced (adaption of microorganisms), and the herbicides were biodegraded, but not below this concentration. Interestingly, after degradation was initiated above 1 µg/L concentrations, degradation continued till concentrations well below the 1 µg/L threshold were reached.

The European Chemical Industry recently initiated a research project which aims at a better understanding of the ecological significance of adaptation.[28] The basis for this was set at a workshop on "Assessing environmental persistence", which concluded, that the importance of biomass concentration and diversity within screening assessments for biodegradability should be better understood, e.g. by considering the adaptation potential due to pre-exposure. Further on, the ecological significance of adaptation should be assessed and appropriate test methods and guidance should be developed (ECETOC 2012).

3.3.3 Test design

3.3.3.1 Reference substances

When considering performing enhanced screening biodegradation tests these should be accompanied by investigating a set of appropriate positive and negative reference compounds which describe the potency of the approach. Comber and Holt (2010) suggested to distinguish between reference compounds which normally pass a ready type biodegradability tests (e.g. aniline, sodium acetate, 1-octanol) and those which normally fail a ready type test but pass an enhanced screening test (e.g. diethylene glycol, 4-chloroaniline, 1,3,5 trimethylbenzene). Further on, considering a negative reference substance, such as di-isotridecyl adipate, terphenyl, or cyclododecane normally failing both, standard and enhanced screening tests, could prevent artefacts in the test design. Martin (2014) recommended the use of non-degradable reference compounds in order to ensure that the test item is not falsely assessed as not being persistent.

3.3.3.2 Endpoints

It has been suggested to retrieve more information from screening biodegradation tests in order to describe more precisely the black box and prevent artefacts in the test design. For example ultimate tests based on oxygen uptake or CO_2 evolution could routinely be combined with additional DOC or parent compound analytics at least at the start and the end of the test, which is already realised in case of the MITI-tests (OECD 301C and OECD 302C). The occurrence of stable transformation products could also be analysed (e.g. Wess and Eisner 2014)[8].

Another option is to describe more precisely the biomass growth during enhanced screening biodegradation tests. Usually, the pass levels of 60% ThOD/ThCO_2 or 70% DOC (for non-adsorbable substances) of the OECD 301 tests assume complete ultimate biodegradation, the remaining part mainly being attributed to biomass growth (see 3.1.4). When the test duration is prolonged, information about actual biomass growth could be used to detect experimental artefacts and prevent false positive assessments. ISO 14852[25] describes a CO_2 evolution test intended for biodegradation testing of plastic materials. The maximum test duration is 6 months.[29] The informative Annex C of this ISO guidance describes an example for the determination of a "cold" carbon balance, by considering CO_2-evolution, DOC in the water phase, biomass growth, and residual polymers. For this, samples of the inoculated medium are taken at the beginning of the test before adding the test material, and at the end of the incubation period. Samples are filtered or centrifuged and the DOC in water is measured next to the biomass (e.g. via protein measurements). The remaining polymers are calculated by weight measurements or specific polymer analytics.

---

[28] CEFIC (2014) LRI-ECO29-Improving assessment of persistency by including adaptation; standardizing methodology and assessing ecological significance. CEFIC Long Research Initiative LRI-ECO29

[29] It should be noted that polymers are excluded from the scope of REACH and that no literature data are available supporting test duration of 6 months.

3.3.3.3 Prolongation of test duration

In the introduction to OECD 301 (1992) paragraph 6 states, that the test normally lasts for 28 days. However, the test duration may be prolonged beyond 28 days when the curve shows that biodegradation has started but that the plateau has not yet been reached by day 28. In this case, the results cannot be used for the classification of a substance as being readily biodegradable. The introduction to OECD testing of degradation of organic chemicals (OECD/OCDE, 2006) suggests (paragraph 21) that only the extent of biodegradation achieved within 28 days should be used for the evaluation of ready biodegradability. However, the degradation extent *after* 28 days would allow the test substance to be classified as *inherently biodegradable*.

That prolongation of test duration can be of utmost importance with respect to certain substances was demonstrated with synthetic cyclohexyl- and norbornyl-derived ketones. In the OECD 301D test, extremely long lag-periods were observed; nonetheless ultimately biodegradation was observed after a test duration of 60 days or even beyond (Seyfried et al., 2015).

For degradability testing of plastic material according to ISO 14852[25] the test duration is completed when the plateau phase has been reached. The maximum test period is 6 months. Longer test durations up to 60 day seem also acceptable for other low water soluble compounds for which the bioavailability is limited and the hydrolysis rate is the bottle-neck for ultimate biodegradation.

For performing enhanced screening degradability tests, often a prolongation of the test duration up to 60 days has been proposed. Especially when a longer lag-phase has been observed and the plateau phase has not yet been reached this seems to be acceptable for persistency evaluation. When the test duration is further extended beyond 60 days the usability of the results for persistency evaluation is questionable. With any prolongation of test duration it is recommended to check the test conditions with additional negative controls in order to detect possible artefacts and also positive test results.

3.3.3.4 Co-metabolism

When a substance is only degraded in the presence of co-substrates which support the growth of microorganisms, this process is called co-metabolism or co-oxidation (Horvath, 1972). In the review document on biodegradability testing Painter (OECD, 1995) proposed to use the definition of Dalton et al. (1982). *"Co-metabolism is the transformation of a non-growth substrate in the obligate presence of a growth substrate or another transformable compound."*

The test design for most standard biodegradation tests consists in incubating the test substance as sole carbon source. One exception is the activated sludge units test for simulating sewage treatment plants in the laboratory, which uses a synthetic sewage as additional substrate. Thus, usually co-metabolism of biodegradation is not taken into account.

Knightes et al. (2006) studied the biodegradation kinetics for binary and complex mixtures of nine polycyclic aromatic hydrocarbons (PAHs) in a non-standardised batch test system. Degradation was followed by chemical analysis. The authors found discrepancies between the observed biodegradation rates and those predicted by a sole-substrate model and concluded that there exist significant substrate interactions which resulted in enhanced biodegradation for all compounds, except naphthalene.

She et al. (2012) found that the biodegradability of 3-nitrophenol, 2,4-dinitophenol and 2,6-dinirophenol improved when glucose was dosed as co-substrate. Similarly, peptone enhanced degradation of crude oil (Mukred et al., 2008).

The promotion of co-metabolism by adding a co-substrate to biodegradation tests has not been supported so far by most authors, although there have been suggestions to use natural instead of synthetic media for improving the possibility of co-metabolism (OECD, 1995).

3.3.3.5 Further proposals for testing difficult substances

Test substances with low water solubility and concomitant low bioavailability can be adsorbed to inert films (often used at Hydrotox laboratory) or silica gel (Handley et al., 2002). Adsorption to inert carriers such as humic acids or silica gel has also been used to reduce toxicity (e.g. van Ginkel et al., 2008).

To enhance bioavailability and hence biodegradation, Rhamnolipids produced by Pseudomonas aeruginosa AT10 were successfully applied to crude oil such that biodegradation extent within 10 days could be increased overall from 32% to 61% and for e.g. isoprenoids from 16% to 70% (Abalos et al., 2004). Similarly, rhamnolipid biosurfactants were investigated regarding their effect on partitioning into the aqueous phase of naphthalene, fluorene, phenanthrene, and pyrene, initially dissolved in di-2-ethylhexyl phthalate (DEHP) or 2,2,4,4,6,8,8-heptamethylnonane (HMN) (Garcia-Junco et al., 2003). However, rhamnolipids were also observed to be toxic to a bacterial strain capable to degrade phenanthrene (Shin et al., 2005). Using non-ionic surfactants (Alfonic 810-60; Novel II 1412-56) at 10 to 100 µg/g with phenanthrene and biphenyl sorbed to aquifer material and soil, desorption was shown to increase. Degradation was significantly increased only in case of competent degrading microorganisms added (Aronstein and Alexander, 1992). Similarly, Chang et al. (Chang et al., 2004) used trehalose lipid biosurfactants at 20-fold their CMC[30] (ca. 320 mg/L) produced by Rhodococcus erythropolis to increase bioavailability of phenanthrene in pure water, soil and a soil-water slurry together with a known degrader for phenanthrene. In effect, rate and extent of biodegradation was enhanced in the water and soil systems, while with regard to the soil-water slurry system only the rate of biodegradation increased, not the extent.

In conclusion, using appropriate agents to increase bioavailability, biodegradability of the bioavailable fraction may possibly be demonstrated to an extent and within a time frame measurable in simple laboratory tests. It may then be assumed that under natural environmental conditions, at least the bioavailable fraction will be prone to degradation while the overall time frame for degradation of the total material (including the sorbed fraction) may be very long. The question remains whether the use of solubility aids adequately simulates the environmental situation.

To simulate natural environmental conditions one can include dissolved organic matter (DOM) in biodegradability tests. From the uptake kinetics of fluorene, phenanthrene, fluoranthene, pyrene, and benzo[e]pyrene by solid-phase microextraction fibres it was shown that the presence of dissolved organic matter (DOM) obtained from sediment pore water increased absorption and desorption rate coefficients (Haftka et al., 2008). Correspondingly, the mineralization of aqueous-phase phenanthrene and pyrene by a competent degrader strain also was found to be enhanced by DOM. Concluding, DOM may have strong implications in processes where adsorption-desorption equilibria are decisive for bioavailability and thus biodegradation.

An interesting approach is a test design enabling primary biodegradability testing of petroleum hydrocarbons in sea water (Concawe, 2012). To provide rapid solubilisation of the complex compound mixtures in seawater in spite of high hydrophobicity and without exceeding respective solubilities in water of constituting compounds, passive dosing was applied: A silicone tubing was used, containing silicone oil saturated with hydrocarbons. Natural seawater was used as microbial inoculum, supplemented with a nutrient solution. Test item analysis was performed by gas chromatography coupled with mass spectrometry. Under these circumstances, most tested compounds had half-lives in seawater of less than 60 days.

---

[30] Critical Micelle Concentration: The concentration for a detergent, from which on surface tension remains relatively constant and aggregates of surfactant will form in solution with increasing concentration.

Similarly, dynamic passive dosing was demonstrated to be feasible in biodegradation tests using a bacterial strain capable of degradation phenanthrene and fluoranthene. Defined dissolved concentrations ranging over 4 orders of magnitude could be tested using this methodology. From these data, first-order mineralization rate constants could be derived. The authors conclude that "dynamic passive dosing avoids using cosolvent for introducing the substrate, buffers substrate depletion so biotransformation is measured within a narrow and defined dissolved concentration range, and enables high compound turnover even at low concentrations to simplify end point measurement" (Smith et al., 2012).

3.3.4 Combinations of several modifications

When performing an enhanced screening biodegradation test, sometimes several modifications are combined. For example, a higher inoculum concentration may be used together with larger test vessels and longer duration. Currently, no guidance exists regarding the limits of these modifications and their combinations, and the conditions under which they still may be acceptable for regulatory persistency evaluation.

To demonstrate biodegradability of the chemical product diaryl-p-phenylene diamine (DAPD) used to inhibit degradation of rubber products, Dailey et al. (2013) combined several enhancements, like elongation of test duration, implication of a radiotracer to enhance analytical sensitivity, enhancement of bioavailability (addition of test item sorbed to silica gel; surfactant) and measurement of mineralization plus assimilation by the microorganisms. The authors report that after 63 days 37% were mineralized. Another 29% was allegedly assimilated or absorbed by the microorganisms.

The transformation of acesulfame in water under environmentally relevant conditions was investigated (Gan et al., 2014), including direct and indirect photolysis, biodegradation, and hydrolysis. While no significant mineralisation was observed, the approach is interesting with regard to a combination of light-induced physicochemical primary degradation events and microbiologically mediated biodegradation.

Separate studies on sulphonamide derivative (sulfamethoxypyridazine) degradation caused by biodegradation on the one hand (not readily biodegradable) and photolysis using a medium pressure Hg-lamp (primary degradation observed) on the other hand are reported (Khaleel et al., 2013). Unfortunately, the combination of both degradation modes was not performed, although this could have been an interesting approach, as transformation products produced by photodegradation could possibly be biodegradable, in contrast to the parent compound.

While the composition of natural inocula is not restricted to bacteria but also contains other organism groups such as algae or protozoa, which also interact in biodegradation processes, the test performance in the dark or diffuse light does not support the growth of algae. Immobilized *Chlorella vulgaris* was applied for degradation of nonylphenol (Gao et al., 2011). However, changes of the light regime in biodegradability tests might also influence the carbon regime (DOC and oxygen production, CO2 consumption) and therefore lower the accuracy of the tests.

A comprehensive review on a set of enhancements for screening tests was provided by Kowalczyk et al. (2015), critically investigating constraints of current screening tests. The main starting point of the authors was that ready biodegradation tests are assumed to produce varying results and may lead to false negative assessments for biodegradability (biodegradable substances not recognised as such and, consecutively, assessed as being persistent). They proposed a number of enhancements such as considering the quality and diversity of the inoculum by microbiological techniques, the use of biofilms as inoculum source, the increase of the test volume for improving the likelihood that rare degraders are contained, and the prolongation of the test duration beyond 28 days, without indicating an upper limit. No combinations of these enhancements, which the authors still considered being

protective, are described in this publication. The authors also refer to the influence of the test concentration of the outcome of biodegradation tests. At higher test concentrations (mg range), as usually applied in ready type tests, the test substance serves as substrate leading to inoculum biomass growth (which is considered in the pass levels of ready type tests). At low concentrations (µg range) as applied in simulation tests, the test substances are regarded to be degraded as secondary (non-growth) substrates concurrently with a variety of natural occurring compounds. On the other hand, higher test concentration my also lead to inhibitory effects.

3.3.5 Interpretation of results

Inoculum blank values represent a validity criterion for the respirometric tests (oxgen uptake or CO_2 evolution), which limits the maximum inoculum concentration and sometimes requires a pre-incubation of the inoculum. These pre-treatments generally lower the biodegradation potency of the inoculum. However, blank values may sometimes not solely be attributable to the inoculum but also the test system: Struijs et al. (1995) analysed the oxygen consumption of activated sludge and secondary effluent from 40 STPs in the two phase closed bottle test and closed bottle test respectively and found a strong dependence on mass loading rate. They also considered additional non-inoculated blank flasks, which were deemed to detect the mineralisation of organic impurities present in the mineral medium. Although the blank oxygen uptake of these non-inoculated flasks was only in the range of 0.2-0.3 mg/L, those values may have a significant influence e.g. on the validity criteria. The authors therefore proposed establishing new criteria for blank values while subtracting the blank values of non-inoculated mineral from the inoculum blank values. This would allow the use of about 20% higher inoculum concentrations without reducing the precision of the method.

It has been suggested to reduce the pass level for OECD BOD tests to 50% (Bealing 2002, Boethling and Lynch 2006, cited after Stasinakis et al., 2008). According to Kowalczyk et al. (2015), there are cases where degradation extents of less than 60% ThOD corresponded with a DOC-elimination above 90%. The authors argued that a pass level of 50% for mineralisation might be more appropriate and has been discussed at OECD, but has not been considered in OECD 301, because of the stringency of the tests. Also, Boethling and Lynch (2007) showed that consistency between results from different screening tests could be improved by lowering the pass criterion to 50% for CO_2 and BOD tests. On the other hand, the differences between mineralisation and DOC-elimination may be explained by adsorption processes (see 3.1.3). Further on, results obtained with readily biodegradable test or reference substances often lead to mineralisation extents far above the pass levels, indicating that the influence of biomass growth often is lower than expected (see also discussion on pass levels when using pre-exposed inoculum in chapter 3.3.2.4).

It is generally accepted that biodegradation screening tests are not designed to predict biodegradation kinetics in environmental compartments (surface water, sediment and soil), due to their unrealistic high test concentration, inoculum concentration, and higher temperature compared to nature. Thus, results from more realistic simulation tests are used to derive biodegradation kinetics, which are used for persistency evaluation (see chapter3.5). The primary aim of performing enhanced biodegradation tests is to potentially fulfil a “pass level” for a substance allowing to predict that it is not persistent under average environmental conditions. For chemicals biodegraded to a predominant portion in enhanced screening tests extensive simulation testing with radiolabelled test materials may thus be avoided. Nevertheless, often kinetic data are derived also from ready type tests or enhanced screening biodegradation tests as additional evaluation of data. While these data may be useful for characterising the degradation kinetics of the substance in screening tests or for comparing the degradation behaviour of different chemicals they should not be used for persistency assessments as such by comparing with half-lives in REACH Annex XIII.

Ahtiainen et al. (2003) compared biodegradation kinetics with the ISO 14593 (1999)[7] headspace CO2 evolution test under environmental conditions. The authors concluded that low concentrations lead to different biodegradation kinetics compared to the high concentrations used in the standard tests. However, with regard to aniline and 4-chloro-aniline, the source of inoculum appeared to have an even higher impact on degradation rates.

Thouand et al (2011) outline how biodegradation success in OECD 301-type tests depends on the cell density of a degradation-competent group of degraders present in the inoculum at time zero. According to the authors, the resulting S-shaped biodegradation curve depended largely on the specific growth rate of these organisms under laboratory conditions as a ratio between food and biomass; and thus, corresponding half-time values derived from those curves would essentially reflect laboratory conditions rather than kinetic characteristics relevant for environmental conditions. In consequence, the authors recommend testing a variety of different inocula and rather derive probability values for degradation than artificial rate constants. The probability for degradation would then reflect the capability of various inoculums (river water, sea water soil, activated sludge etc.) to degrade a substance under realistic exposure conditions and food/biomass ratios.

Federle et al. (1997) compared experimental data for nine chemicals under carefully controlled screening and simulation tests: Ready biodegradability (OECD 301 B, CO2 evolution) – ^{14}C batch activated sludge test at 1 mg/L test concentration – ^{14}C river water test at 2 mg/L – ^{14}C soil mineralisation test at 1 mg/kg test concentration. For the latter 3 tests, radioactive material had been used. All nine chemicals were mineralised in the test systems, but a clear relationship of the biodegradation kinetics between the screening test (OECD 301 B) and the simulation tests could not be established. The authors concluded that data from ready tests cannot be used to predict biodegradation kinetics in real environmental compartments.

The question remains whether a positive result obtained from (enhanced) screening tests can predict that a substance is not persistent in simulation tests or the environment. In a research project on behalf of the German Environment Agency (Moltmann and Gartiser, 2001) available screening test data have been compared with data from the water/sediment and soil simulation tests. In the end, 113 pairs of screening and simulation tests (mainly pesticides) were available. Only 8 active substances were found to be "mineralizable" (pass level 60% ThOD/$ThCO_2$ reached) in the screening tests, 4 of which showed a better rating in the screening test than in the simulation test. On the other hand, 34 substances were better degradable in the simulation test than in the screening tests. Differing results between both test categories could partly be explained by the test conditions or performance. The results demonstrate that screening tests usually do not tend to predict biodegradability or non-persistency too favourable.[31]

3.4 Solid phase screening tests

3.4.1 Water-sediment screening tests

A shake-flask test was proposed by Cripe et al. (1987) for determining the biodegradability of organic compounds at low concentration levels (200 µg/L) with and without suspended sediment (500 mg/L). Natural water and sediment was used instead of e.g. sewage sludge inoculum. Since non-radiolabelled compounds were used, test substance concentrations were determined by electron capture

[31] It should be noted that most data were derived from pesticides because at that time no data requirements existed for other substance groups. Meanwhile, further data e.g. from REACH, the Biocidal Product Regulation or the Environmental risk assessment of pharmaceuticals are available, which would also be worth being evaluated.

gas-liquid chromatography. The method enables to determine first order rate constants and associated half-lives in the range of 1 to 30 days. It was shown to be suitable to describe natural adaption phenomena and – by parallel incubation of sterile samples and solvent extraction to discern between degradation and adsorption. Sediment enhanced biodegradation is described for methyl parathion. Further, to account for the draw-back of the test to observe primary biodegradation, only (by monitoring of parent), in parallel acute toxicity tests were performed with mysids or daphnids at time zero and during the degradation phase to integrate ecotoxicity of possibly stable transformation products (decline of parent should normally be associated with a decline of ecotoxicity).

Flenner et al. (1991) investigated the effect of sorption to suspended sediment on the degradation of different n-alkyl esters of p-aminobenzoic acid, covering a large range of hydrophobicity. As a result, suspended sediment caused a reduction of the overall biodegradation rate that was rapid initially but then slowed down, which was attributed to the sorbed fraction.

Parsons (1992) examined the influence of suspended sediment on the biodegradation of chlorinated dibenzo-*p*-dioxins. After seven days, the sorbed fractions were sufficiently readily desorbed to be partly degraded. However, biodegradation rates were lower in the sediment suspensions than in the solutions without sediment.

A laboratory sediment water system was used to study sorption, isomerization and biodegradation potential of hexachlorocyclohexane under aerobic and anaerobic conditions (Wu et al., 1997). Addition of organic nutrients (glucose mixed with yeast powder) dramatically increased isomerization from α- to β-HCH and biodegradation especially under aerobic conditions.

Xia and co-workers investigated the effect of various levels of sediment on biodegradation of polycyclic aromatic hydrocarbons (PAH) in natural river water. Results show that biodegradation rates of PAHs increased with the sediment content. This is explained by desorption of sorbed PAHs at the water-sediment interface, where also most of the bacteria are present, resulting in an increased contact between bacteria and PAHs (Xia and Wang, 2008; Xia et al., 2006; Xia et al., 2011). Sediment enhanced biodegradation was also observed for methyl parathion by Cripe et al. (1987) in the shake-flask test described above and was preliminarily attributed by the authors to greater microbial biomass associated with sediment particles.

Water-sediment water-soil screening tests were applied to study the entwined processes of sorption and biodegradation, using pharmaceuticals acetaminophen, caffeine, propranolol, and acebutolol, differing pronouncedly in their adsorption coefficients (Lin et al., 2010). It could be demonstrated that removal of all tested compounds was possible by natural attenuation (dilution, hydrolysis, photolysis, biodegradation, dispersion, irreversible sorption; generally reducing the toxicity of contaminants towards the environment and human populations) and that suspended sediments can significantly affect their fate and behaviour in the aquatic environment.

A water-sediment screening tool (WSST) implying an artificial sediment layer protected from turbulence by a fine meshwork was recently designed based on OECD 301C methodology (Junker et al., 2010). MITI inoculum is used, involving pre-culturing and starvation immediately before the test. Prove of principle investigation demonstrated the test protocol to be feasible and applicable, in spite of a higher background compared to OECD 301C. From comparison of test results for aniline and benzoic acid from this test with the original OECD 301C (without sediment), the authors conclude that a lowering of the pass level for the sediment version should be considered (50% rather than 60%).

The WSST was further applied to determine experimental mineralization rates and kinetics for fifteen organic chemicals by means of non-linear regression models (Junker et al., 2016). The experimental results showed good reproducibility and in most instances where in the same range as degradation data from well-established methods (e.g. OECD 308) found in literature and databases. The authors

therefore conclude that the WSST can be used to determine sound and reliable quantitative mineralization data including mineralization kinetics on the screening test level in addition to the water-only OECD 301C.

A screening water sediment test (WST) was published by Baginska et al. (2015) implying an artificial and standardized medium based on existing OECD guidelines and the work of Junker et al. (2010). Artificial sediment was adjusted to improve the oxygen penetration into the sediment by reducing the clay content from 20% to 5%. Furthermore, the peat content was lowered from 2% to 1% in order to reduce the background respiration of the sediment. To limit the decrease in bacterial diversity preculturing of the mixed inoculum was shortened from one month to ten days. For optimization and validation, aniline, diethylene glycol and sodium acetate were used, the applicability was tested with two pharmaceuticals, acetaminophen and ciprofloxacin. The authors conclude that the water sediment test proved to be a promising tool for the biodegradation investigation of chemicals at the water-sediment interface.

3.4.2 Soil screening tests

For the development of soil screening tests it is certainly important to consider differential behaviour with regard of degradation half-lives: depending on whether low concentrations of isotope-labelled substrates are used or rather high concentrations of non-labelled compounds for respirometric measurements, different biodegradation kinetics are expected. This was compared for the biodegradation of three different surfactants in soils and sludge-soil mixtures by use of 14C-labelled compounds and automated respirometry (Gejlsbjerg et al., 2003). An increase in concentration from 10 mg/kg to 400 mg/kg caused a reduction in relative maximum mineralization rate and an accompanying increase in lag time of approximately a factor of 3.5.

A soil screening tool (SST) has currently been published by Junker et al. (2016). Field-fresh standard soil Lufa Type 2.3 adjusted to a moisture content of 45% of the water holding capacity was used to investigate the mineralization of fifteen organic compounds in parallel to the original OECD 301C (MITI-Test) and the water-sediment screening tool (WSST; Junker et al., 2010). Results could be verified by showing good agreement with soil mineralization data (e.g. from OECD 307) for the same compounds in literature and databases for most of the test compounds.

3.4.3 Interpretation of results

The need to incorporate the compartments soil and sediment in biodegradation tests was already expressed in the early 1980s (Van der Harst et al., 1981). In the following, several methods for testing the sorption and biodegradability of organic compounds in the presence of sediment have been published. However, their suitability as a screening test is largely limited, since low concentration levels and radiolabeled test compounds are used, resulting in high costs for chemical analyses and procurement (e.g. Ingerslev and Nyholm, 2000; Jensen et al., 1988; Xia et al., 2011). The requirement for screening biodegradability in water-sediment systems was renewed recently (ECETOC, 2013c), but existing methods for ready biodegradability do not consider the effect of sediment or soil on degradation.

Beyond that, theoretical models (e.g. QSARs) to predict the environmental fate of compounds were mainly developed based on qualitative biodegradation data related to water-only test systems since semi-quantitative biodegradation data from soil and sediment are often lacking.

In general, the conclusions drawn regarding the interpretation of results from standard screening tests for ready biodegradability (see 3.1.4) are also valid for solid phase screening tests. Thus, it is generally assumed that a compound meeting the criteria for ready biodegradability in screening tests will also undergo rapid and ultimate biodegradation in the environment and consequently no further investigation of the biodegradability is required. If the chemical fails the ready biodegradability test it

might nevertheless be degraded under relevant environmental conditions and further higher level testing (e.g. a test on inherent biodegradability or a simulation test) is required.

However, some characteristics have to be taken into account, e.g. with regard to validity criteria, relating to the application of solid matrices like sediment and soil. Since the publication by Junker at al. (2016) is to our best knowledge the first one describing the application and evaluation of water-sediment screening tests and soil screening tests in parallel to a standard ready biodegradability test, the following paragraphs are mainly based on this work.

Processes like sorption, ageing and sequestration may affect the degradation of a compound. Sediment or soil contains organic matter (e.g. peat), which plays an important role regarding sorption and has inter alia been proposed to be the most significant factor for organic compound interactions with sediments (Cornelissen et al., 1998). Junker et al. (2016) described the application of experimental screening tools for water-sediment and soil in addition to the water-only test according to guideline OECD 301C. Substance-specific differences were observed for mineralization among the three test systems. However, the observed differences do not reflect the (reversible) sorption into organic matter in terms of K_{OC} values.

According to OECD 301 the pass level of 60% mineralization should be reached for the reference compound by day 14. This pass levels could not be reached for aniline in water-sediment screening tests and for sodium benzoate in soil screening tests (Junker et al., 2016). Consequently, the authors recommend a pass level of 50% mineralization at day 14 as validity criterion for aniline in the water-sediment system and for sodium benzoate in the soil system. Aniline did not appear suitable as reference compound in soil, since mineralization was too low.

On the other hand, peat might be used by the microorganisms as source of carbon and energy and thus may increase the background respiration. This is confirmed by Junker et al. (2016) who measured an oxygen consumption of 30.2 ± 13.4 mg O_2/L (mean ± standard deviation after 28 days) for OECD 301C, whereas clearly higher background respiration was observed in the water-sediment screening tool (79.9 ± 18.6 mg O_2/L) and in the soil screening tool (734.0 ± 186.9 mg O_2/L). Based on the 95% percentiles, the authors therefore recommend values of 110 mg O_2/L for the water-sediment system and 1100 mg O_2/L for the soil system to be used as validity criterion for oxygen uptake of the inoculum blank. An option to lower the background respiration of the artificial sediment is to reduce the peat content as suggested by Baginska et al. (2015).

Although the use of natural sediments and associated water in screening tests would be desirable in terms of environmental relevance of the test system, this would result in higher variability due to different characteristics of the sediment matrix depending on location and season of sampling. Moreover, the addition of complexity into test systems may affect the outcome of biodegradation tests in a manner which is hard to predict (ECETOC, 2013c). Therefore, standardized media should be used for solid phase screening tests in view of reproducibility and comparability of results. For soil, however, sources of supply exist for field-fresh standard soils (e.g. Lufa, Speyer, Germany) at a constant quality and composition. Thus, soil screening tests can be performed by making use of natural soil microorganisms. It has to be mentioned that screening tests cannot consider the multitude of existing soil characteristics (e.g. particle size distribution, pH, organic matter content, microbial community), which might affect biodegradation.

Solid phase screening tests provide the opportunity to investigate the biodegradation of organic compounds in the presence of the compartment of concern (the compartment the substance is likely to partition to) and in consideration of processes that affect bioavailability and biodegradation (e.g. sorption, desorption). Recent developments (e.g. Baginska et al., 2015; Junker et al., 2016; Junker et al., 2010) describe promising methods to examine the biodegradability of compounds in sediment and soil on the screening test level. However, further research (e.g. more compounds to be tested in

comparison to standard screening tests) is needed to put their applicability on a firmer scientific footing and to establish prediction models to unambiguously identify substances as degradable in these screening tests.

3.5 Simulation tests

3.5.1 Overview

The OECD introduction to biodegradability testing (OECD/OCDE, 2006) states that chemicals that fail to meet the criteria for ready biodegradability or even inherent biodegradability may be rapidly degradable when present at low concentrations in the environment. Simulation tests may be used to examine the biodegradation of organic chemicals in STPs, soil, aquatic sediment, and surface water. If it can be demonstrated that the chemical is ultimately degraded by more than 70% in 28 days under realistic conditions in the aquatic environment (e.g. by using OECD 308 or 309), then the definition of "rapid degradability" in relation to aquatic hazard classification is met. The results of a simulation test may show a rapid transformation of the parent compound, whereas ultimate degradation (measured by e.g. CO_2 gas production) is limited due to the formation of recalcitrant transformation products. It is therefore necessary to distinguish between primary and ultimate biodegradation, when the rate and extent of degradation are calculated. Whenever possible, assessment of biodegradation in the environment should be based on results from tests simulating the conditions in the relevant environmental compartment. Man-made organic chemicals will normally be present at low concentrations (i.e. low µg/L level) in the environment compared to the total mass of biodegradable carbon substrates. This implies that the anticipated biodegradation kinetics are first order ("non-growth" kinetics). If a higher concentration is used in a test (e.g. to examine transformation products), biodegradation of the chemical will frequently support growth of the degrading microorganisms (OECD/OCDE, 2006, paragraph 50-56).

The challenges in "Determining real-world biodegradation rates" are shortly summarized in an older editorial by Howard (1993). Fifty percent disappearance times for methylene chloride varied over 100-fold with different soils, and half-lives of linear alcohol ethoxylate varied "only" 10-fold within 11 different soils. A 10-fold variation means 20 vs 200 days, the latter value being already far above the threshold for persistence in soil of 120 day according to Annex XIII. This demonstrates the considerable impact of soil characteristics on the degradability of a compound.

Ericson (2010) compared the results of the CO_2 evolution test (OECD 301B) and the OECD 314B test carried out with 5 pharmaceutical active substances. While Eplerenone, Atorvastatin, Varenicline, and Sunitinib malate showed relative low degradation extents in both test systems (0-23%), the biodegradation of Exemestane in the OECD 314B reached 81% compared to only 15% in the OECD 301B. The author attributed the test results obtained in the OECD 314B as proof of ready biodegradability (which is not the correct interpretation when considering the test design as simulation test.)

Berkner and Thierbach (2014) collected OECD 301 and 308 (water sediment simulation) test data on active ingredients of pharmaceuticals. They compared overall mineralization extents between both test systems as well as the data available on accumulating metabolites from the simulation test. In nearly 45% of the studies, half-lives for transformation products were longer than those for parent compounds. The overall mineralisation extent was usually low according to both tests. In an attempt to investigate the impact of test conditions on degradation behaviour, water-sediment simulation tests with three pharmaceuticals (log Koc values of 3.6, 4.0 and 4.7) were performed and evaluated (Ericson, 2007), with an assessment of aerobic and anaerobic degradation as well as of non-extractable residues (using radiotracers). In conclusion, up to 94% non-extractable residues were found, an-

aerobic conditions resulted in less biotransformation and mineralization compared to aerobic conditions, and authors recommend performing the test under anaerobic conditions only for those compounds amenable to typical anaerobic processes. The most comprehensive review of results from OECD 308 tests was performed by Ericson et al. (2014). Total system half-lives for 31 different pharmaceuticals are evaluated, considering their physico-chemical properties, non-extractable residues as well as transformation products. Concluding from these results, recommendations for improvement of the test as well as for a more consistent and transparent interpretation of the results are given, e.g.: design of a more relevant water-sediment transformation test reflecting the typical discharge scenario for humane pharmaceuticals (larger water : sediment ratio); consistent use of terminology and consistent interpretation of results; use of parent compound first order total system half-lives ($DegT_{50\text{-system}}$) instead of disappearance times for water and sediment, respectively; research on cationic pharmaceuticals: whether their classification as such might already be sufficient as an alert to high levels of non-extractable residues (NER); the issue of bioavailability of these residues (Ericson et al., 2014).

Honti and Fenner (2015) analysed 41 experimental OECD 308 data sets for pharmaceuticals and pesticides. They found that disappearance half-lives (DT_{50}) can easily be derived but they lump degradation and phase transfer information and are not robust against changes in test system geometry. Degradation half-lives ($DegT_{50}$) are less system-specific, but require inverse modelling, resulting in considerable uncertainty. The results support concerns about the usability and efficiency of the experiments. Nevertheless, the authors suggest the $DegT_{50,system}$ as a useful indicator of persistence in the upper aerobic sediment layer. However, the test system geometry should be reported.

As an alternative to Phase II Tier A testing of pharmaceuticals, Ericson (2010) suggested to apply the OECD 314B method (adopted 2008; OECD, 2016) for activated sludge. This method consists of an open or sealed batch or flow-through system. The test substance is incubated for e.g. 28 days at environmental relevant concentrations (µg/L range) with high concentrations of activated sludge (2500-4000 mg d.s. /L). The closed flow-through system is preferred, when the $^{14}CO_2$ produced should be trapped for determining the level of mineralisation.

OECD simulation testing guideline 309 (aerobic mineralisation in surface water) allows for some possible modifications. For example, the test can be conducted as a "pelagic test" with surface water only or as a "suspended sediment test" with sediment amounts between 0.01 and 1 g/L. However, there is a lack of experience with this test system and the potential impact of modifications on degradation is largely unclear. Völkel and Höger (2015)[32] evaluated 24 tests according to OECD 309. They conclude that depending on the chosen study design, this "standard" test can be both very simple and very complex. They also stress that the effect of suspended solids on the degradation rate is still unclear.

In a case study on degradation of the fungicide isopyrazam according to OECD 309, the influence on degradation rate of diffuse light as well as inclusion of inoculum of suspended sediment was investigated. Diffuse light was used at an intensity representative of deeper layers of large, open water bodies (<7% of the incident intensity; light-dark cycle), and it was demonstrated that metabolism of isopyrazam by phototrophic microorganisms was rapid, whereas degradation in continuous darkness was negligible. Different light intensities resulted in similar degradation rates (DT_{50} 38 days at 7% of incident light intensity compared to 48 days at only 2% of incident light intensity), indicating that sufficient light for photosynthesis to occur is necessary while there seemed no direct proportional ef-

[32] Völkel, W.; Höger, S.J. (2015): Aerobic Mineralisation in Surface Water (OECD 309): Experiences and Interpretation. Poster presentation SETAC Europe Barcelona 2015

fect of light intensity on biodegradation rate. In contrast, inclusion of an inoculum of suspended sediment did not have a large impact on degradation (Hand and Moreland, 2014). This emphasizes the importance to include daylight equivalent light-dark cycles into biodegradability testing, also within screening tests, for integrating the degradative potential of phototrophic organisms. Within OECD 309 (surface water simulation test), "diffuse light" may be used as an alternative set-up instead of incubation in the dark, to ensure survival and activity of phototrophic microorganisms that are present in the system, whereas direct or indirect photolysis of the test substance should be precluded as far as possible.

Within the Cefic-funded project LRI-ECO18 (http://cefic-lri.org) on "Identifying limitations of the OCED water-sediment test (OECD 308) and developing suitable alternatives to assess persistence", a suite of four different water/sediment systems was used to investigate the behavior of four reference substances with varying sorption properties and biodegradability in two different natural sediments in order to bridge the gap between the OECD 308 and 309 tests: (1) the OECD 308 standard protocol (water/sediment ratio = 3:1); (2) a modified OECD 308 protocol (water/sediment ratio = 10:1, stirred water phase); (3) a modified OECD 309 protocol (water/sediment ratio = 100:1, stirred system), and (4) an OECD 309 standard protocol (water/sediment ratio = 1000:1, stirred system). Beyond that, Bayesian parameter estimation and system representations of various complexities were used to evaluate existing OECD 308 data and to estimate degradation rate constants from individual experiments as well as combinations of experiments.

The results of the project as well as experiences with OECD 308 and OECD 309 have been discussed recently at an ECETOC workshop (6 October 2015, Dübendorf, Switzerland). The following findings, recommendations and research needs regarding test performance, derivation and interpretation of degradation half-lives were presented and will be published soon:

Experiments

- Thinner sediment layer and stirred water phase of modified OECD 308 resulted in a thicker oxic sediment layer.
- Mineralization was increased in modified test systems but mostly went hand in hand with increased formation of non-extractable residues (NER). Formation and assessment of non-extractable residues (NER) need further investigation. This topic is addressed within two Cefic-funded projects (LRI-ECO24 and LRI-ECO25, see http://cefic-lri.org/).
- No differences could be observed between the four test systems regarding ^{14}C mass balances and variation of test results.
- OECD 309 test systems should be shaken rather than stirred to prevent confounding processes, e.g. increased sorption/NER formation during the experiment due to sediment grinding.
- The current OECD 309 guideline allows too much variability of the experimental setup, which has a significant influence on results. Thus, the guideline needs a review and relevant parameters should be harmonized.

Derivation of degradation half-lives

- Derivation of compartment-specific degradation half-lives for sediment ($DegT_{50,sed}$) and water ($DegT_{50,w}$) from OECD 308 data alone is highly uncertain and not recommended.
- Dissipation half-lives for water and sediment ($DT_{50,w}$ and $DT_{50,sed}$) are confounded by phase transfer processes and should not be used for comparison to persistence criteria or for exposure modeling.

- The total system degradation half-life ($DegT_{50,system}$) is to some extent system-dependent. System geometries (e.g. inner diameter of test vessels, heights of water and sediment column, sediment dry weight used) should be fixed or at least fully reported.

Interpretation of degradation half-lives

- For substances with high sorption potential (K_p > 2000 L/kg) $DegT_{50,system}$ can be considered a good surrogate indicator for persistence in sediment.
- For less strongly sorbing substances, a bioavailability-corrected and biomass normalized second-order biodegradation rate constant k'_{bio} can be derived if both OECD 308 and 309 data are available. This value can be converted back to a half-life for sediment, which can be used for persistence assessment. In addition k'_{bio} can be considered to be used as a compartment-independent indicator of biodegradation potential. The conceptual soundness and the applicability of k'_{bio} have to be further validated with additional data sets.
- Substances that are not readily biodegradable are likely to fulfil the persistence criterion for water unless they are hydrolyzed. An evaluation of existing OECD 309 data against the persistence criterion for water is needed to consolidate this finding.

Simplified "simulation tests" for surface water and sediment based on either radiolabelled substrates or specific test item analytics were proposed by Ingerslev and Nyholm (2000). The work aimed at establishing a simple shake-flask surface water biodegradability die away test using environmentally relevant test item concentrations between 1 and 100 µg/L, thus meant to provide information on biodegradation behaviour and kinetic rates. "Used with surface water alone the test simulates a pelagic environment and amended with sediments (0.1-1 dry weight/L) the test is intended to simulate a water environment with suspended solids (e.g., resuspended sediments)." The test was further examined with regard to lag times depending on the test volume used, and increased random failure was observed for small sample volumes (Ingerslev et al., 2000). Finally, the test system was applied to 7 example compounds ranked for their relative biodegradability: aniline>p-nitrophenol, 2, 4-dichlorophenoxyacetic acid>4-chloroaniline>maleic hydrazide, pentachlorophenol>atrazine. The test was applied for determining first-order rate constants for the primary biodegradation of four antibiotics applied at intermediate concentrations (50-5000 µg/l) (Ingerslev et al., 2001). A comparison to biodegradation rate constants for these same compounds determined using simulation tests according to OECD is not given.

Using 0.37 µg kg^{-1} [14C]-4-nitrophenol, a laboratory simulation test with natural water and sediments was evaluated as a means to study mineralization of chemicals present at low concentrations in surface waters (Kalsch et al., 1999). The effects of different important parameters were evaluated, including sediment type, time of sediment collection, aeration methodology, illumination and temperature. Besides mineralization of 4-nitrophenol, the distribution of radioactivity between the different compartments and the physicochemical and biological state of the sediment–water systems were studied. Finally, considering in addition the results of experiments with lindane, a test guideline for standardised testing is supposed by the authors. This work preceded adoption of OECD 308 in 2002.

The STP simulation test according to OECD 303A and OECD 314 is a specific case because its design differs from all other simulation tests. First, the test concentration is relatively high (10-20 mg/L DOC), thus allowing substantial growth of competent degraders. ^{14}C labelled test substances are rarely used and thus no carbon balance is usually established. Second, the synthetic sewage dosed to the system allows for co-metabolism processes. Third, the OECD 303 intends simulating STP which are rather technical than natural compartments. Thus, the transferability to other environmental compartments such as surface water, sediment or soil is rather limited and test results should not be used for the P-assessment of these compartments.

3.5.2 Interpretation of results

Simulation studies according to OECD guidelines 307 (soil), 308 (aquatic sediment) and 309 (surface water) are an integral part of the tiered testing strategies for the environmental risk assessment of chemicals including persistence assessment. Before testing, the compartment of concern has to be identified considering uses and release patterns as well as physical-chemical properties of the test substance. In REACH guidance R.11 (ECHA, 2014c) a flow-diagram illustrates how to select the compartment of concern and consequential the appropriate simulation test.

The purpose of the tests is to measure the time-dependent degradation of a (^{14}C-labeled) test substance at environmentally relevant low concentration levels (e.g. from 1 µg/L to 100 µg/L in OECD 309) in order to ensure "non-growth" biodegradation kinetics. The rate and route of degradation of the parent compound and, if possible, its transformation products will be followed throughout the test period of 60, 100 or 120 days for surface water (OECD 309), water-sediment (OECD 308) and soil (OECD 307), respectively. A complete ^{14}C-mass balance for each sampling time point will be established including non-extractable residues (NER). The endpoints usually derived from simulation studies are primary and ultimate degradation rates and half-lives ($DegT_{50}$) or dissipation half-lives (DT_{50}) for the compartments included in the test system. However, within the persistence assessment only degradation half-lives ($DegT_{50}$) should be compared to the persistence criteria of REACH Annex XIII (ECHA, 2014c; Rauert et al., 2014).

Simulation tests have been extensively used within different regulatory frameworks to derive persistence indicators since the respective guidelines were adopted. However, there are still open questions, particularly with regard to evaluation and interpretation of simulation tests according to OECD 308 and OECD 309. This topic will be discussed in chapter 5.5.

Non-extractable residues (NER) are often formed in considerable amounts in simulation studies with soil (OECD 307) and sediment (OECD 308, OECD 309). However, there is much debate on how to define, how to determine (i.e. which extraction methods should be used) and how to interpret (i.e. if they are bioavailable or might become bioavailable in the future) NER in the regulatory context. Different positions exist, ranging from NERs considered as an efficient toxicant removal process, to NERs interpreted as a sink and thus a potential future source of toxicants (ECETOC, 2013b). An extraction methodology framework has been developed to distinguish between 'bioavailable' and 'bioaccessible' residues (ECETOC, 2013a; b). Definitions for non-extractable residues (NER) and bound residues (BR) are given in ECETOC special report 18 (ECETOC, 2014). It is concluded that "NERs are strongly bound to sediments and while adsorbed they are protected from degradation." Furthermore, a statement is made that "although these NERs remain in the environment they are not bioavailable and therefore in the context of PBT assessment they should be considered equivalent to not being 'P' or 'vP'." However, studies on aged lysimeter cores with the herbicide atrazine (Jablonowski et al., 2009) have shown that the unchanged substance can be remobilized from those residues under environmentally relevant conditions even after many years of ageing, which results in essentially the same concern as for persistent compounds. Hence, an unconditional exclusion of non-extractable residue formation from persistence assessments may not be warranted.

Another upcoming issue is the test temperature. Most simulation studies have been performed at a temperature of 20 ± 2°C, which is appropriate according to the simulation test guidelines. However, a temperature of 12°C is considered as representative mean temperature of European surface waters. Thus, a temperature correction of existing degradation half-lives might be necessary. Moreover, testing at lower temperatures representative of the climatic conditions (e.g. 12°C) is often required by authorities. Although this is comprehensible with regard to environmentally relevant conditions, the experimental implementation in the laboratory is often challenging and expensive

4 Regulatory requirements for identification of persistent compounds

4.1 Persistence according to REACH: Annex XIII – Identification of PBT and vPvB substances

Substances persistent in the environment are of concern as they may accumulate in a given environmental compartment, if there is a recurring influx caused by direct or indirect exposure and the substance of concern is of sufficient immobility (no or low efflux). Then, over time concentrations may reach levels that are harmful for humans and the environment. This is even more the case for compounds with bioaccumulative properties. Their presence in the environment at low levels may suffice for accumulation of toxic levels in organisms over time.

According to article 1 paragraph 3, REACH "... it is for manufacturers, importers and downstream users to ensure that they manufacture, place on the market or use such substances that do not adversely affect human health or the environment." In doing so, the precautionary principle has to be applied. This general provision of REACH certainly does also apply to the PBT assessment according to REACH Annex XIII: it is therefore up to the manufacturer or importer of a substance to conclusively demonstrate that P properties (and also B, or T properties) are not fulfilled. Otherwise the compound is suspected to be persistent and has to be treated as if it would fulfil P criteria.

To identify those compounds, criteria for persistence are specified in REACH Annex XIII. The actual criteria are degradation half-lives for environmental compartments, as outlined in Table 1. Half-lives ($DegT_{50}$, for *degradation time 50*) are meant to be first order or pseudo-first order half-lives, and thus independent from concentration. According to REACH guidance document R.11 (ECHA, 2014c) it is not appropriate to compare 50% disappearance times with the criteria given in REACH Annex XIII.

Table 1: Degradation half-lives for identification of PBT and vPvB substances (according to REACH Annex XIII)

Environment	Terrestrial, Fresh water & Estuarine		Marine	
Compartment	**P**	**vP**	**P**	**vP**
Water	40 days	60 days	60 days	60 days
Sediment	120 days	180 days	180 days	180 days
Soil	120 days	180 days	--	--

These criteria have to be understood as disjunction, i.e. fulfilment for one compartment suffices to qualify a compound as persistent (P) or very persistent (vP).

Tests resulting in half-lives directly comparable to these values (termed *assessment information* in Annex XIII, see Table 2) are restricted to so-called *simulation tests* for the different environmental media (e.g. OECD 307, 308, and 309) and *other information*, e.g. field studies and monitoring studies. While not explicitly mentioned, tests on hydrolysis (e.g. OECD 111, "Hydrolysis as a function of pH") do also provide first-order rate constants for degradation, and thus are one further example for *other information*. However, hydrolysis demonstrates per se only primary degradation, implicating the need to assess hydrolysis products for possible PBT or vPvB properties. Similarly, in simulation tests, where full mineralization could not be demonstrated and rather degradation of the parent substance is followed by substance specific analysis (i.e. primary biodegradation), resulting degradation products must be identified and assessed for PBT properties. Follow-up on primary degradation products is an explicit requirement of Annex XIII: "... relevant constituents of a substance and relevant transformation and/or degradation products ..." have also to be assessed for PBT/vPvB properties. With

regard to hydrolysis rate constants, according to guidance document R.11 (ECHA, 2014c), rate constants measured in pure water may not reflect rate constants in sediment or soil, especially for compounds prone to adsorption, such that partitioning and a potential for ionisation must be taken into account here. Therefore, fast hydrolysis rates alone cannot lead to a conclusion of non-persistence.

Table 2: Assessment and screening information on P as specified in REACH Annex XIII

Information category	Type of applicable tests	Type of information gained
Assessment information on P	(a) Results from simulation testing on degradation in surface water; (b) Results from simulation testing on degradation in soil; (c) Results from simulation testing on degradation in sediment; (d) Other information, such as information from field studies or monitoring studies, provided that its suitability and reliability can be reasonably demonstrated.	Half-lives in different environmental media ($DegT_{50}$) – directly comparable to criteria given in Annex XIII (see Table 1)
Screening information on P	(a) Results from tests on ready biodegradation in accordance with Section 9.2.1.1 of Annex VII; (b) Results from other screening tests (e.g. enhanced ready test, tests on inherent biodegradability); (c) Results obtained from biodegradation (Q)SAR models in accordance with Section 1.3 of Annex XI; (d) Other information provided that its suitability and reliability can be reasonable demonstrated.	Generally, the output from these tests is qualitative only, and based on the respective pass levels, results are yes- / no-type with regard to biodegradability. Thus, in case pass levels are fulfilled, "reasonable" biodegradability in environmental media, including the marine environment, is assumed and the compound regarded to be not persistent (not P).

According to Annex XIII these data should be evaluated in a "weight of evidence determination using expert judgement" considering all available information (including *screening information*) and weighing the data by quality and consistency. *Assessment information* used shall have been "obtained under relevant conditions" (see Table 2 for details). This weight of evidence assessment includes screening information as well as data from simulation tests (section 3.2 of Annex XIII). Most probably, this is due to the equivocal results often obtained from simulation tests: $DegT_{50}$-values may depend heavily on the exact specimen of medium (e.g. soil, sediment, surface water) used (for some examples on the extent of variation possible, see e.g. Howard, 1993). This may be due to differences in the degree of adsorption caused by different constituents of the particular matrix (e.g. clay minerals, metal oxides) and associated parameters like pH, cation exchange capacity or redox potential; repercussions of these parameters on biodegradation; as well as differences in the microbial density

and microbial diversity contained in different specimens of one media type. With regard to adsorption and similar effects often subsumed under this term, non-extractable residues (NER, for more details see section 5.6.3) are often an issue. However, current REACH guidance does not offer any help here, neither with respect to how one should deal with non-extractable residues nor regarding as to how one should interpret differences in half-lives found for different media specimens. REACH Guidance on Information Requirements and Chemical Safety Assessment Chapter R.11: PBT/vPvB assessment (ECHA, 2014c) more specifically deals with the use and applicability of different types of existing tests for persistence assessment, while REACH Guidance on Information Requirements and Chemical Safety Assessment Chapter R.7b: Endpoint specific guidance (ECHA, 2014b) gives the background on biodegradability and interpretation as well as use of existing guideline tests with respect to REACH).

Highly volatile compounds (high Henry constant), for which photolysis in air may be the most important degradation pathway, are not considered in the regulation: No cut-offs for half-lives in air are specified in Annex XIII. According to REACH guidance document R.11 (ECHA, 2014c) a degradation half-life in air > 2 days indicates a potential for long-range atmospheric transport with reference to the Stockholm convention on POPs; and accordingly, such compounds might be transported to and deposited in remote areas. The same value for half-life in air is suggested by Scheringer et al. (2006), when discussing persistence criteria for compounds air. Concluding, a lack of degradability in the atmosphere is currently assessed for long-range transport potential rather than persistence. The difficulty of biodegradability testing of highly volatile compounds is further outlined in section 3.1.5.3).

4.2 Information sources for P-evaluation

4.2.1 Simulation tests

As outlined in section 4.1, cut-off values for half-lives as given in REACH Annex XIII generally can be obtained from simulation tests (but also from other so called *assessment information*, "provided that its suitability and reliability can be reasonably demonstrated"). The half-lives according to REACH Annex XIII are meant as pseudo-first order half-lives (see section 4.1), i.e. simulation-type tests must be designed such that equivalent results may be derived. However, according to Rauert et al. (2014) "best fit kinetics such as first-order multi-compartment (FOMC) can be used if they are recalculated by dividing the DegT90 values by a factor of 3.32 or by using the degradation rate constant of the slower phase in case of double first-order in parallel model (DFOP) or Hockey stick (HS) kinetics."

While simulation type tests are the basic information source according to Annex XIII criteria, a weight of evidence approach is demanded within this Annex as outlined above, including situations where simulation test data are available. When testing biodegradability in different environmental compartments, a compound observed to fulfil P criteria in at least one environmental compartment must be considered to be persistent.

More difficult to assess are contradicting results for one compartment from several tests. In this context, the notion within Annex XIII that results shall have been "obtained under relevant conditions" could be interpreted in a way that a half-life obtained using e.g. a soil more representative for either the European conditions or the relevant exposure situations may be given a higher weight compared to (equally reliable) results from a soil less representative. This is supported by the notion within OECD 307 (aerobic and anaerobic transformation in soil), that "the types of soils tested should be representative of the environmental conditions where use or release will occur."

How the variability of environmental conditions is to be implemented into simulation type tests is poorly defined. While at least four different soils have to be assessed according to OECD 307 for the scope of deriving transformation rates, only two different sediments have to be assessed according to

OECD 308 (sediment simulation test); and according to OECD 309 (water simulation test) one single surface water is sufficient to derive rate constants for degradation in surface water. Similar to soils, sediments may be very different in composition; and also for surface waters differences in degradation potential (e.g. dependent on eutrophication, discharges from sewage treatment plants) cannot be excluded. Guidance on how varying results from simulation tests with different specimens should be interpreted is lacking.

With regard to evaluation of simulation test results, Rauert et al. (2014) give some suggestions how multiple results for one environmental compartment could be evaluated, depending on the amount and quality (reliability) of the data. Beyond others, the authors suggest in case of few (up to four) studies for one compartment (same quality and reliability) to select the worst-case $DegT_{50}$. With 5 and more reliable values available, taking the geometric mean is suggested, or – more conservative and taking into account the range of variability – taking the 90th percentile. Where data are not equally reliable, data should be weighted according to their quality and the range and distribution of all values assed by a weight of evidence approach.

With regard to the water-sediment simulation test (OECD 308), REACH guidance document R.7b (ECHA, 2014b) states: "Although for substances with Kp >2000 [ca. log K_{OC} 4.3] an aquatic sediment simulation test might be relevant in addition to a pelagic simulation test, a good test of this type **does not exist** yet." 'And further with regard to uncertainty considerations: "...it is uncertain what the value of conducting the strict anaerobic test part of the OECD 308 test is, and how these data can be used in CSA." Rauert et al. (2014) outline in this respect, that no standardised tests are currently available measuring true degradation in the sediment compartment. OECD 308 is used to assess the fate in water and sediment system, but reliable separate $DegT_{50}$ values for water and sediment cannot usually be derived from the study results. The authors suggest comparing $DegT_{50}$ for the total system to both, trigger values for water as well as sediment. In case of compounds with an equilibrium pronouncedly on one side (sediment or water), DegT50 could be compared to the respective trigger value alone.

4.2.2 Information from (enhanced) screening tests

By applying conservative conditions screening tests aim at identifying substances which are with high certainty **not** persistent. REACH guidance document R.11 (ECHA, 2014c) lines out that "it is normally not possible to conclude [whether] the substance fulfils the PBT or vPvB criteria due to the uncertainties related to screening information." However, some exemptions are discussed in the guidance document:

- "If test results are available showing that a substance is not inherently biodegradable under the mentioned conditions [specific criteria] this is a clear indication that the substance will not biodegrade in the marine environment and, hence, must be regarded as persistent."
- "Lack of degradation (<20% degradation) in an inherent biodegradability test equivalent to the OECD TG 302 series would provide sufficient information to confirm persistence without the need for further simulation testing."

Further, conclusions on *not P* from existing screening tests must follow the criteria as specified by REACH guidance documents R.11 (ECHA, 2014c) and R.7B (ECHA, 2014b). Thus, substances

- assessed to be readily biodegradable according to OECD 301/OECD 310 testing guidelines or equivalent, irrespectively of the 10 day window requirement (see section 3.1.4 for details on cut-offs)
- fulfilling pass levels for ready biodegradability in enhanced biodegradation screening tests (see below)

- being inherently biodegradable according to OECD 302 B and C testing guidelines *and* fulfilling specific criteria:
    1. 302 B (Zahn-Wellens test): ≥70% mineralization (DOC removal) within 7 d; log phase no longer than 3d; removal before degradation occurs must be below 15%; pre-adapted inoculum must not be used.
    2. 302 C (MITI II test): ≥70% mineralisation (O_2 uptake) within 14 days; log phase no longer than 3d; pre-adapted inoculum must not be used.

can be considered being neither vP nor P. However, if these criteria are not fulfilled, and in the absence of higher tier data (simulation tests), persistence is *assumed* ("potentially P or vP"), but actually no final conclusion can be drawn from screening tests, as outlined by guidance R.11.

This interpretation of test results of *inherent biodegradability tests* according to REACH guidance may be considered conservative compared to the view of OECD (OECD/OCDE, 2006). This document outlines that "... inherent biodegradability can be considered to be a specific property of a chemical [and] it is [therefore] not necessary to define limits on test duration or biodegradation rates." While the intention of REACH Annex XIII goes significantly beyond this in that biodegradation in the respective media must proceed in a reasonable time frame and further, translation of screening test results to environmental compartments must be conservative, these *specific criteria* were originally set within EU-TGD (EC, 1996) as a prerequisite to assign generic rate constants for environmental exposure assessment. These criteria were meant to assure that "the elimination in the test can really be ascribed to biodegradation, and; no recalcitrant metabolites are formed, and; the adaption time in the test is limited." Taking this into account, the 7-day limit to achieve the pass level in the Zahn-Wellens test seems quite conservative with regard to evaluating persistence.

OECD (OECD/OCDE, 2006) further explains that "When results of ready biodegradability tests indicate that the pass level criterion is almost fulfilled (i.e. ThOD or DOC slightly below 60% or 70%, respectively) such results can be used to indicate inherent biodegradability. This is also the case when the pass level criterion is fulfilled but the 10-day window criterion is not." While this is exactly stated this way within REACH guidance document R.7b (ECHA, 2014b), this interpretation of inherent biodegradability is of no use in persistence assessment according to REACH guidance document R.11 (ECHA, 2014c), where Table R.11-4 explicitly specifies that ready biodegradability test results below the respective pass levels of OECD 301 are to be taken as a trigger for classifying a substance as "potentially P or vP", thus leaving no room for other interpretations in case of results slightly below these levels.

Enhanced biodegradation screening tests are based on tests on ready biodegradability (OECD 301, 310); they include, however, modifications deviating from the guideline, which are aiming at achieving a higher probability for degradation as compared with the relatively stringent ready test designs (for details, see section 3.3). Accordingly, pass levels with regard to CO_2 production, O_2 consumption (>60%, each) and DOC removal (>70%) are essentially equivalent to those defined in the corresponding OECD guideline documents. Possible modifications are outlined within the REACH guidance document R.7b (ECHA, 2014b) and are described in more detail in section 3.3. These tests may only be used to demonstrate non-persistence; rate constants to characterise biodegradation under environmental conditions cannot be derived from these tests. However, if the following statement of guidance document R.7b is to be taken seriously, the practical relevance of enhanced biodegradation screening tests would be low: "The enhanced screening tests are restricted to using only natural environmental media as the source of inoculum e.g. marine and freshwater. Enhanced screening studies using inocula derived from sewage treatment works **cannot** be used in persistence assessments." This will be further discussed in chapter 5.

According to REACH guidance document R.11, pass levels with regard to CO_2 production, O_2 consumption (>60%, each) and DOC removal (>70%) within 28 days are also applicable to standardized *marine biodegradability tests* (OECD TG 306, Marine CO_2 Evolution test, Marine BODIS test, and the Marine CO_2 Headspace test; see sections 3.1.1 and 3.1.3 for details). However, as outlined in OECD TG 306 and confirmed in REACH guidance document R.7b (ECHA, 2014b), the maximum admissible incubation period is up to 60 days (Shake Flask Method), (a longer incubation period is considered reasonable due to the slower degradation in sea water). Owing to technical constraints, generally 28 days apply for the marine variant of the Closed Bottle Method. Further, "When a chemical attains >60% ThOD or >70% DOC removal in a Biodegradability in Seawater test (OECD 306), it can also be expected to fulfil the criteria for ready biodegradability" (ECHA, 2014b).

Apart from accepted simulation and screening tests, Annex XIII is open for considering "other information". This applies to both, screening as well as assessment information on P provided that "suitability and reliability" can be reasonably demonstrated. Assessment information shall have been "obtained under relevant conditions" (see section 4.1), because only then reliable first order rate constants can be derived and half-lives compared to the cut-off values given in Annex XIII.

With regard to new screening type tests ("other information"), these must be shown to be suitable and reliable. Interpreting this, reproducibility should be warranted, and results should be sufficiently conservative to yield meaningful prediction of *not P*. Currently available screening tests (tests on ready biodegradability, inherent biodegradability or enhanced screening tests) do *not intend* to simulate any environmental compartment – therefore, test design must be stringent and the prediction model used for interpretation of test results (yes/no) must be sufficiently conservative to cover also environmental conditions unfavourable for biodegradation.

REACH guidance R.11 (ECHA, 2014c) outlines that "although it might be theoretically possible to calculate degradation half-life values from screening information, such values cannot be directly compared with the P/vP criteria of Annex XIII to REACH, but the screening information should be discussed as such and compared with the screening criteria [as given within guidance document R.11]." All screening tests (including newly developed enhanced or compartment-type tests) therefore should have prediction models leading to (biodegradation) yes/no results. Nevertheless, rate constants may be determined in such tests and help to interpret results and compare between test results within one test type. But interpretation of rate constants should be confined to the conditions of the screening test.

In addition, results from ready biodegradability tests may be used for assessment of biodegradation in a specific environmental compartment if no data from simulation tests are available (OECD/OCDE, 2006; Rauert et al., 2014). REACH guidance R.7b (ECHA, 2014b) states that "it is not always necessary to know the exact degradation half-live value but rather simply that it is above or below the threshold."

Several methods to derive rate constants or half-lives from extent of biodegradation in screening tests have been described in literature (e.g. Aronson et al., 2006; EC, 2003; Jaworska et al., 2003; Junker et al., 2016) and are used within QSAR models (e.g. US EPA EPI Suite[33] and PBT Profiler[34]).

---

[33] US EPA, United States Environmental Protection Agency (2012). Estimation Programs Interface Suite™ (EPI Suite) for Microsoft Windows, v 4.11. United States Environmental Protection Agency, Washington, DC, USA. Available at: http://www.epa.gov/opptintr/exposure/pubs/episuitedl.htm

[34] US EPA, United States Environmental Protection Agency (2012). PBT Profiler. Version 2.000, September 4, 2012. United States Environmental Protection Agency. Available at: http://www.pbtprofiler.net.

Thus, screening information including rate constants or half-lives derived from media-specific screening tests (water, water-sediment, soil) might be considered to make judgements within a persistence screening or as part of a weight-of-evidence-based persistence assessment. However, they should not be used for deriving an unequivocal conclusion on persistence within a definitive assessment.

4.3 Conclusions

With regard to P assessment, conclusive results on persistence are desired already at the screening level, as simulation type tests are performed only under certain circumstances (B, T fulfilled or cannot be excluded; environmental exposure assessment indicates the need) and thus rarely available. Ideally, (new) screening type tests should

- be significantly faster and cheaper than simulation tests
- be reproducible and reliable
- produce reliable outcome with regard to the diverse conditions encountered in the environment (ideally no false negatives for P; ideally a low number of false positives for P)
- enable predictions for various compartments.

It is obvious from the discussions above that these expectations are difficult to reconcile in one test. Screening tests are fulfilling three out of the four conditions above, but currently often err to the conservative side. Simulation tests may produce reliable results for the specific conditions of the test, but fail to be fast and cheap and do not allow predictions for other compartments than the one tested.

Chapter 5 will discuss potential modifications and improvements of various types of tests.

5 Data gaps for P assessment and deficits of existing tests

5.1 Screening tests for ready biodegradability

5.1.1 Deficits and possibilities to improve test performance

In the analysis of existing testing approaches the following deficits in terms of accuracy, repeatability, and comparability have been identified for the ready type screening tests:

- The inoculum concentration of the OECD 301 ready biodegradability tests is mainly limited by the expected values of the inoculum blanks (validity criteria for the blank values), but the existing borderlines of different tests are not consistent (see chapter 3.1.1). A systematic definition of upper inoculum blank values is lacking. Based on the percentage of the maximum theoretically possible CO_2 production or theoretical O_2 consumption due to the amount of test compound present in the test, currently valid upper blank values for OECD 301 tests range from ca. 21% in the closed bottle test[35] up to 55% for the CO_2 evolution test.
- The inoculum used in biodegradation tests should better be described e.g. in terms of the MLR of the activated sludge used or additional microbiologic parameters (see chapter 3.3.2.1).
- Attempts to reduce the inoculum blank by pre-incubation and/or pre-treatment may allow higher inoculum concentrations but also have their drawbacks, because often the inoculum potency is reduced (whereupon the extent of such reduction may even differ for different types of degraders within the medium) (see chapter 3.3.2.2).
- Many of the studies performed to demonstrate the influence of the inoculum source, quality, pre-treatment, concentration, and total amount have been carried out under test conditions not comparable to standard OECD 301 tests (low volume tests in microplates, use of activated sludge supernatant instead of dry solids etc.). Thus, the consequences for performing ready type (or enhanced) screening tests for persistency evaluation still remain unclear.
- The number of replicate vessels in OECD 301 tests is considered as being too low and should be increased for enhanced biodegradability testing, including inoculum blanks (e.g. the MITI-I test requires only one inoculum blank replicate). Further on, criteria for the variability of the inoculum blank values in parallel vessels should be established. The current OECD 301 methods only prescribe a maximum allowed variability of 20% of the degradation extent in parallel vessels, but do not indicate an allowed variance of the inoculum blanks.
- Potential adsorption to activated sludge should be carefully examined when using DOC based tests such as OECD 301A. No guidance on what adsorption extent may be acceptable exists. The elimination through adsorption should be limited by defining a clear criterion e.g. of a maximum of 20% DOC-elimination through adsorption at the test start being acceptable. Further on, the consideration of abiotic controls (test item with inorganic toxic substance but without inoculum) considerably improves the identification of adsorbable or volatile test substances.
- The water-only test systems do insufficiently consider processes like sorption and desorption which may affect bioavailability and degradation.

[35] assuming 7 mg/L ThOD for test item concentrations of 2-10 mg/L

5.1.2 Possibilities to improve their applicability for P assessment

The possible improvements for the ready biodegradation test described in chapter 5.1.1 may also help their applicability in persistency assessment. A positive result in a ready type test can be assumed as criterion for non-persistency, when the pass levels of 60% ThOD/$ThCO_2$ or 70% DOC-elimination have been reached within 28 days irrespective of the fulfilment of the 10-day-window. When considering biodegradability testing of poorly soluble test items several improvements have been described (see section 3.1.5.2). In this case the use of (additional) reference compounds of poor solubility but ready biodegradability could improve the predictability of the test design. Microcrystalline cellulose could be an example of a suitable reference compound, which usually is ultimately biodegradable at least in some OECD 301 tests, while not fulfilling the 10-day window. Validated and accepted reference compounds to check these methods for difficult substances are lacking.

The OECD 301 allows the use of several types of (mixed) inocula, the maximum inoculum density being limited by the validity criterion established for the inoculum blank (e.g. OECD 301 B: 40 mg /L CO_2). Laboratories could improve the potency of their tests by using inoculum densities which come up near to the upper limit allowed for the inoculum blanks or by using increased flask volumes without questioning the ready type test approach. Further modification for performing enhanced biodegradability tests are described below.

5.2 Enhanced screening tests for ready biodegradability

5.2.1 Currently discussed modifications

Enhanced screening tests are not designed for determining ready biodegradability of a test substance but exclusively to allow an evaluation as not being persistent. The intention is to avoid the performance of extensive simulation tests. Among the modification are the prolongation of the test duration beyond 28 days, the use of larger test vessels, the increasing the biomass concentration, and pre-exposure to the test substance at low concentrations (see chapter 3.3 and 4.2.2). The inoculum source and quality is considered being of major importance. The increase of the test vessels as well as of the inoculum concentration mainly intends improving the possibility that competent degraders are contained. Usually only test results fulfilling the pass levels of 60% ThOD /$ThCO_2$ or 70% without the application of the 10-day-window are acceptable for assessing a substance as not being persistent. The derivation of kinetic degradation rates is not intended because their transferability to natural compartments is not given.

According to REACH Annex XIII, information to be used for persistence assessment includes screening type information with explicit reference to “enhanced ready test(s)” as well as other information, provided that data are suitable and reliable. REACH guidance document R.7B (ECHA, 2014b) demands that only inoculum ***not*** derived from STP could be used for enhanced biodegradation screening tests to demonstrate non-persistence. As most ready-type tests are performed with STP-derived inocula, and one of the most often applied enhancements is a prolongation of incubation time beyond 28 days, this restriction would significantly restrict applicability of this test type.

5.2.2 Test designs and conditions suitable for use in P assessments

So far precise guidance on enhanced biodegradation tests is lacking. Rather, indications are given on possible enhancements (ECHA, 2014b), occasionally without specifying exact upper limits and without giving details on possible combinations of enhancements. Thus, there is a need for critical assessments of what modifications are considered being acceptable to yield suitable and reliable information needed for a screening-assessment of persistence according to Annex XIII:

a) In REACH guidance on PBT/vPvB assessment the term "enhanced screening test" is not clearly defined although sometimes the term "enhanced ready biodegradation test "is used (ECHA, 2014c). Thus, it should be clearly stated that enhanced screening tests are restricted to ready type tests and "enhanced inherent screening tests" do not exist. As a consequence the maximum inoculum concentration allowed in enhanced screening test should be below that used for inherent type tests in order to maintain their nature of being ready biodegradability tests.
b) The restriction of performing enhanced screening tests with inocula **not** derived STP is considered not being meaningful for the following reasons:
First, activated sludge and secondary treated effluents are the inocula most often applied in the ready type screening tests according to OECD 301 and OECD 310 which are considered being stringent. Because these inocula are allowed for ready type biodegradability tests and these tests are accepted to be relevant for the "Non-P"-assessment, this is not consistent with the REACH guidance with regard to enhanced screening tests.
Secondly, inocula derived from STP are not completely out of natural sources. The contribution of STP effluents to the river flow in central Europe is assumed to be 10%. Thus, a dilution factor of 10 is usually considered for STP effluents in the exposure assessment according to the ECHA guidance R.16 (ECHA, 2012). Gartiser (1999) used statistical effluent and river flow data for estimating the fluctuation of wastewater ratios in German rivers. While this ratio considerably differs depending on the river basin and the extreme flows (low flow or flood flow) the mean ratio of STP effluents in Germany approximates 10%, which is well in line with the default value of the guidance document. Similarly, effluents from STP also contain suspended solids derived from activated sludge. Certain nutrients such as phosphorous or nitrogen mainly are emitted from STPs in covalently bound or adsorbed form. This is the reason why the European Urban Wastewater Directive 91/271/EEC has restricted the total suspended solids in STP effluents to a maximum of 35 mg/L which is in the same range as the inoculum density in ready type tests.
Third, inocula derived from STP seem to be more reproducible than e.g. surface water or filtered soil eluates, which highly depend on local or seasonal variations.
Thus, inocula derived from STP should be considered being acceptable for enhanced ready biodegradability testing as far as no pre-adaptation to the test item is envisaged.
It is emphasized here that neither ready biodegradability tests nor enhanced ready tests claim to represent real environmental conditions.
c) The use of inocula from contaminated sites or pre-exposed to low test item concentrations is considered not being suitable for a prospective persistency assessment and should therefore be excluded.
d) The prolongation of the test duration is not considered being critical because these modifications are already mentioned in OECD and REACH guidance when the plateau phase has not been reached. Positive results obtained from prolonged tests should not be used for ready biodegradability but only for P assessment. For practical reasons a upper limit of 60 days could be defined as suggested in the REACH guidance (see chapter 3.3.3.3).
e) The increase of the test vessels is not considered being critical because the OECD 301 only gives an indication of a suitable size. The upper limit of the test vessel size will depend on the practicability of the test device in terms of air tightness or stirring capacity (see chapter 3.3.2.3).
f) The consideration of *co-metabolism based degradation* is currently not considered in the design of enhanced screening tests. However, this is an important mechanism of degradation for certain compounds (see chapter 3.3.3.4). The option for including a co-substrate in enhanced screening tests is not very promising because the introduction of a new biodegradable carbon source may cause a considerably additional uncertainty in tests based on unspecific sum parameters such as DOC-elimination, oxygen depletion, or CO_2 evolution.
g) When performing enhanced screening tests additional positive and negative substances should be considered, which reflect the complexity of the test system. For example, when a prolongation

of the test duration is intended a reference compound which is also degradable only at the longer term should be used. Positive reference compounds normally biodegradable under enhanced but not under ready type conditions as well as negative controls normally not biodegradable also under enhanced conditions should be considered in the test design (see chapter 3.3.3.1).

h) The pass levels for enhanced screening tests to be used for persistency evaluation have still not been defined. As a first approach the same pass levels for P/not P equal to that specified within the OECD 301 series should be applied. But there is also an option to define other pass-level as a consequence of certain enhancements (significant prolongation of degradation time frame could theoretically result in a higher assimilation rate (i.e. incorporation in biomass compared to catabolism to CO_2).

i) No validity criteria specific for enhanced biodegradation screening tests exist. The criteria given in OECD 301 often are not applicable anymore when certain enhancements are applied. The main issue is that the differences between measurements of the test and blank vessels should allow accurate estimates of the degradation extent. This difference mainly depends on the ratio of the test concentration compared to the inoculum concentration. The higher the inoculum activity the lower is the accuracy of the measurements if the test concentration is maintained (which also may be limited for inhibitory substances). The consideration of additional replicate vessels for both the test and blank vessels allow a more precise estimate of the variability of the measurements and therefore of the accuracy of the biodegradation extent.

The question remains, which enhancement as such and what combinations of enhancements would still be acceptable from the regulatory perspective leading to results which allow conclusions on degradation behaviour (and thereof persistency) under environmental conditions. As a general principle, all combinations of optimizations of ready biodegradation tests which are not questioning their category as "ready type test" should be allowed. Examples are the increase of flask volume, the use of mixed inoculum within the validity criteria, the use of additional positive and negative controls or the increase of replicate vessels. However, the combination of enhancements beyond the test design of ready biodegradability tests, such as an increase of the inoculum density, needs to be critically assessed. Although screening tests are not intended to represent environmental conditions, the use of results from such test systems in the environmental risk assessment should require that the potential for extrapolating from artificial test conditions to environmental conditions is appropriately validated when test designs are modified or enhanced. It is suggested here

- to define clear limits of single enhancements (as discussed above, e.g. inoculum concentration below that of inherent tests, no acceptance of pre-exposed inoculum, maximum test duration 60 days, test vessel size as appropriate)
- to evaluate the impact of combining several test modifications and/or enhancements and to exclude combinations which impair the "screening" nature of the enhanced tests
- and to critically discuss performance of such enhanced ready tests based on results for substances with known biodegradability. Based on results with these reference substances agreement on suitable pass levels and validity criteria should be achieved.

5.3 Screening tests for inherent biodegradability

5.3.1 Deficits and possibilities to improve test performance

Inherent tests are performed under more favourable conditions and thus give useful information whether any potential for biodegradation exists irrespective of their relevance to environmental compartments. Results from inherent biodegradation tests may be used for assessing persistency in two ways. First, test results above 70% are used for indicating ultimately biodegradability and are used as trigger for non-persistency when specific criteria (lag phase no longer than 3 days, pass level

reached within 7 days) are met. Second, negative results from inherent tests (<20% DOC-elimination) may indicate the potential for persistency (ECHA, 2014b). Blok et al. (1984) concluded from their simulations of the influence of the quality and quantity of the inoculum, that only a negative result obtained in the Zahn-Wellens test (OECD 302 B) has a strong predictive value for assessing a substance as being persistent while a positive result has a low discriminative value.

The following problems and possibilities for improvements with inherent tests were identified:

- There is no standardized inherent CO_2 evolution test existing, although in literature several methods based on the OECD 301B and the OECD 310 have been developed The combination of two endpoints, CO_2 evolution and DOC-elimination in one test has successfully been applied and would more precisely allow distinguishing between mineralisation and adsorption.
- The MITI II test (OECD 302C) requires to use a very specific pre-incubated inoculum mixture, but other inoculum sources such as activated sludge seem also acceptable (Beek, 2000)[36]. Guidance on acceptable inoculum sources for the MITI II test is missing. From a regulatory view the use of other inocula such as activated sludge (OECD 302 B) or mixed inocula derived from different environmental compartments (activated sludge, lake and river water) may be acceptable.
- Potential adsorption to activated sludge should be carefully examined when using DOC based tests such as OECD 302B. The three hours value in the Zahn-Wellens test should be indicated in all test reports - not only when there is suspicion for adsorption (see chapter 3.2).

5.3.2 Test designs and conditions suitable for use in P assessments

The applicability of inherent biodegradability tests for persistency evaluation is restricted because only tests fulfilling the specific criteria described in chapter 4.2.2 allow an assessment as not being persistent. Because of the high inoculum concentration used for these tests, which is representative only for STPs, this approach seems being justified. On the other hand, this implies that he upper limit of the inoculum concentration applicable for enhanced screening tests should be well below that used for inherent type tests (e.g. < 200 mg/L d.s. activated sludge)(see chapter 3.3.2.3).

5.4 Media-specific screening tests

5.4.1 Currently discussed test designs and their deficits

Media-specific screening tests comprise processes which may affect bioavailability and degradation (e.g. sorption) and thus can provide compartment-specific information on degradation. Such tests allowing for derivation of environmentally relevant half-lives (apparent first order) would be closest to simulation tests with regard to media-specific assessment of persistence. However, taking into account the high test item concentrations used therein that are needed for achieving good signal to noise ratios without using radiolabelled test items hampers the determination of environmentally relevant kinetics from those tests.

Nonetheless, some published but non-standardized media-specific screening tests were developed (see section 3.4). It can be assumed that these tests will yield results closer to reality for those media (suspended matter, sediment, soil) as compared to results of screening tests using sewage sludge without further addition of solids:

---

[36] Chapter *The Assessment of Biodegradation and Persistence* (Beek, B., Böhling, S., Franke, C., Jöhncke, U., Studinger, G., Thumm, E.)

- adsorption processes are mimicked, at least to a certain extent, depending on the exact setup of the test;
- correspondingly, degradation rates are expected to be lower in cases where adsorption is an issue;
- degradation rates will be more media-specific if such a media-specific inoculum is used (which may or may not be the case, depending on the exact test system).

To achieve acceptability of those tests for P assessment it may therefore be necessary

- to adapt maximum admissible incubation times for reaching respective pass levels for O_2 consumption or CO_2 production to values above 28 days (see e.g. 60-days limit for the screening test for degradation in marine water according to OECD 306)
- to adapt valid pass levels for O_2 consumption or CO_2 production to values below the currently foreseen 60%-level (see e.g. Junker et al. (2010) recommending 50% instead of 60% based on their initial results with a sediment-water system).

Up to now no in-depth evaluation has been undertaken and no guidance exists within REACH guidance documents regarding these relatively new screening-test types.

5.4.2 Test designs and conditions suitable for use in P assessments

Media-specific screening tests (e.g. for sediment and soil) can be used to indicate rapid biodegradability in the same way as tests for ready biodegradability. Readily biodegradable compounds are considered to be "not P" whereas compounds failing the criteria are rated as "potentially P" (see section 4.2.2).

Although several media-specific screening tests for water-sediment and soil have been developed, there is still no standardized test system available that is generally accepted and can be applied for persistence assessment.

To establish media-specific screening test systems that can be used for regulatory purposes, the following criteria should ideally be fulfilled:

- A consolidated test design should be agreed that provides reproducible and sound biodegradation data, suitable to identify (potentially) persistent compounds. This could be achieved by investigating the experimental method in an inter-laboratory comparison (ring test).
- Taking into consideration the screening type nature of these tests prediction models need to be developed with clear criteria, which lead to decisions of being (potentially) persistent or not. Such prediction models might use degradation pass levels with or without specific degradation rates or other test outcomes but will ultimately lead to a yes/no conclusion on persistence. In general, direct comparison with Annex XIII criteria is not recommended, as test conditions do not reflect environmental conditions.
- The resulting test designs and prediction models then should be applied to determine biodegradation data for a defined set of test compounds (e.g. reference compounds identified by Comber and Holt (2010) to check modified and new biodegradability test methods.
- The results should finally be compared to results for the same compounds obtained from standardized screening test according to OECD 301 and – if available – to biodegradation data from literature and databases from well-established simulation studies for the respective compartment.

5.5 Simulation tests

5.5.1 Deficits and gaps of existing tests for use in P assessments

Several deficits and gaps of existing simulation tests have been identified and discussed over the past years. Although some issues are also related to OECD 307 on "Aerobic and Anaerobic Transformation in Soil", e.g. the question how to determine and interpret non-extractable residues (NER), the main focus of discussions was on limitations of OECD 308 and OECD 309.

Even if a test according to OECD 308 is conducted correctly, evaluation and interpretation of results can be difficult (Solomon et al., 2013). The main points of criticism can be summarized as follows, based on ECETOC (2010), Radke and Maier (2014) and the outcomes of the Cefic-funded project LRI-ECO18 (http://cefic-lri.org):

- The continuous exchange between water and sediment in the test system is neither standardized nor quantifiable. Therefore, it is not possible to determine robust compartment-specific degradation half-lives.
- The redox conditions in the sediment are undefined, since the test simulates an aerobic water column over a thin aerobic sediment layer followed by a deeper anaerobic layer. Consequently, the test results are influenced by factors that affect the oxygen distribution in the sediment, e.g. aeration rate, turbulence in the water phase, sediment depth and texture.
- The recommended water:sediment ratio of 3:1 to 4:1 (v/v) together with the recommended sediment height of 2.5 ± 0.5 cm does not represent 'natural' conditions. Due to the small ratio equilibrium mass distribution is often shifted towards the sediment phase.
- The resulting range of system geometries (e.g. height of water and sediment layer, interfacial area) can influence distribution processes (being governed by the partitioning equilibrium constant and the diffusion rate from the water/sediment interface into bulk sediment) as well as biodegradation processes and thus affect persistence in the experimental system. Hence, results are to some extent an artefact of the test system, which hampers the transferability to environmental conditions.
- The static design of the test also does not represent natural conditions by not accounting for the effects of flow velocity and sediment dynamics.

Beyond that, there is a lack of experience with the surface water simulation test (OECD 309) and the various options with regard to the test design complicate the interpretation of results as well as a comparison of studies, e.g. with or without suspended sediment; with or without diffuse lighting.

In addition, the following comprehensive topics are still subjects of debate:

- How to define, determine and interpret non-extractable residues (NER) and bound residues (BR) within persistence assessment.
- Which temperature should be used in laboratory simulation tests (e.g. 20°C or 12°C).

5.5.2 Possibilities to improve test performance and evaluation

Several recommendations how to improve test performance and evaluation, particularly for OECD 308 and OECD 309 have been made in literature and have been presented recently at the ECETOC workshop on "Identifying limitations of the OECD water-sediment test (OECD 308) and developing suitable alternatives to assess persistence" (6 October 2015, Dübendorf, Switzerland).

- OECD 309 test systems should be shaken rather than stirred to prevent increased sorption during the experiment due to sediment grinding.
- The wide array of experimental options in OECD 309 leads to very different outcomes. The test setup should therefore be specified.

- Within OECD 309, the option to include "diffuse light" must be elaborated: clear specifications on light-dark cycles should be given; intensity, duration and wave length distribution must be defined.
- Test system geometries should be fixed or at least fully reported, particularly for OECD 308, since the total system degradation half-life ($DegT_{50,system}$) is to some extent system-dependent.
- New modelling approaches might allow the prediction of compartment-independent biodegradation parameters (e.g. the bioavailability-corrected and biomass-normalized second-order biodegradation rate constant k'_{bio}) that can be used to assess persistence at the water-sediment interface.

Furthermore, to improve the applicability of simulation test results for persistence assessment, clear guidance is needed against which persistence criterion (water, sediment or a new criterion) total system half-lives ($DegT_{50,system}$) should be compared as well as how to deal with multiple different results for one compartment.

5.6 Special considerations for substances difficult to test

5.6.1 Poorly water-soluble substances

Currently, poorly water soluble substances[37] in principle can be assessed for degradability using screening type tests, especially if specific modifications are introduced to account for this difficult substance class (see sections 3.1.5.2 and 3.3.3.5). However, compared to soluble compounds there is an imminent risk of underestimating biodegradability, out of the following reasons:

- For most compounds, only the soluble fraction is accessible for biodegradation.
- Even for compounds which can be degraded in the solid state, the accessible surface area is more or less restricted, depending on the practically possible way of presenting the material within the test. Therefore, the fraction of test item effectively available for biodegradation may be considerably lower than the nominal concentration, leading to slower degradation.
- In screening type tests, biomass growth is a consequence of the high substrate concentrations generally present in these tests. Biomass growth determines the observed sigmoidal biodegradation curve. Biomass growth will be much slower if the available amount of substance (soluble fraction or accessible surface area) is limited due to poor solubility. This effect may be less pronounced if either the available surface area is high (small particles) for compound accessible in the solid state; or both of the following is given: solubilisation is not kinetically hindered and equilibrium is established fast, such that the effectively available concentration will approximately equal the solubility limit; the solubility limit concentration still promotes bacterial growth in a sufficient manner.

As such, if the effectively available concentration is low in screening type tests, in extreme cases there may be no relevant growth of biomass and biodegradation may be severely underestimated. In contrast, due to the much lower concentration present in simulation tests, the nominal concentration applied may be close or below the solubility limit even for poorly soluble compounds and effective degradation may be observed (if the compound is not per se recalcitrant to biodegradation).

---

[37] There is no uniform definition of poor water solubility available: REACH guidance R.7b refers to OECD (< 100 mg/L) and EU-TGD (2003) (< 1 mg/L), outlining that rather below 1 mg/L problems are expected to occur with regard to environmental fate and ecotoxicity testing,

5.6.2 UVCB and multi-constituent substances

A systematic and practically feasible concept for the assessment of UVCB compounds is lacking: while an assessment of single constituents down to 0.1% w/w is demanded, the requirements for specification of the UVCB substance only requires compounds ≥ 10% to be specified by IUPAC name and preferably a CAS number (ECHA, 2014c). Constituents are often poorly defined, rendering radioactive labelling practically impossible, as well as application of QSAR models. Thus, while screening tests may be the only practicable way of demonstrating degradability, REACH guidance restricts their use to the case of UVCB consisting of homologous structures. In this case, if > 60% mineralisation based on CO_2 production or O_2 consumption within 28 days can be demonstrated, it can be assumed that the substance is not persistent in the environment. In case of UVCBs based on structurally non-homologous constituents, biodegradability "should be judged on a case-by-case basis" based on "relative composition and degradability of individual constituents" which often may not be possible for this type of substance. Further, it is suggested for UVCBs to assign constituents to groups with similar properties (a block – in analogy to the hydrocarbon block method) and assess each group for degradability, but the guidance document is lacking any further practical details here.

In principle, multi-constituent substances (MCS) pose a similar problem: The different single constituents may degrade to a different degree and / or at a different rate. As outlined above for UVCB, grouping is recommended and assessment of biodegradability group by group might be an option. Other than for UVCBs, constituents of MCS are defined, such that QSAR or a READ-ACROSS approach to characteristic substances for each group with known biodegradability might be possible.

5.6.3 Substances adsorbing to or reacting with matrices - simulation tests with sediment and soil

One frequent and relevant challenge in the interpretation of simulation studies in soil and sediment is the formation of non-extractable residues (NER). Sediment and (even more pronounced) soil is a complex meshwork of inorganic and organic matter, providing a multitude of different residues for interaction: hydrophobic interaction, binding by van der Waals forces, charge-transfer complexes, polar, ionic or covalent interaction may be relevant, alone or in combination or sequentially as a function of time (e.g. *ageing process*, where hydrophobic compounds slowly sorb to soil organic matter and become increasingly recalcitrant to extraction). Entrapment into soil matrix pores eventually also leads to theoretically reversible, but practically irreversible (due to a very slow release) binding to soil (Kästner et al., 2014). This non-extractable fraction of a compound present at the end of a biodegradation simulation test with sediment or soil is difficult to evaluate. The extent of NER formation can be remarkable. Approximately 12% of pesticides mentioned in a review with data collated for 97 compounds were found to have a proportion of NER formation larger than 70% with regard to mass balance of radioactive material (Barriuso et al., 2008). The nature of NER is mostly not assessed by routine experiments and extraction procedures applied as a means to quantify non-extractability. Moreover, methodology for extraction and characterization of bound material is far from being standardised or generally agreed (Kästner et al., 2014; UBA, 2015). Based on isotope mass balancing, NER may be composed of xenobiotic NER (parent or metabolites) bound to the soil matrix either by non-covalent or covalent interaction; or biogenic NER, where the isotope label is assimilated in biomolecules. While the latter is regarded to be of no concern, both, parent or metabolites may be remobilized even when covalently bound and therefore may not be regarded as degraded (Kästner et al., 2014).

Summarising, for substances forming non-extractable residues a valid experimental procedure and associated accepted conclusions as to when a substance is no longer substance but becomes "matrix" is lacking (see sections 3.5 and 4.1). Recently, compilations of extraction procedures and assessment of their applicability became available (e.g. ECETOC, 2013a; UBA, 2015), but the conclusion is that

considerable further research and standardisation is needed here. A scientifically based agreement is needed on how to

- experimentally assess non-extractability: currently applied organic solvent based extraction is certainly applicable for non-charged, hydrophobic compounds; however, this may not be the case for e.g. positively charged compounds bound by ionic interactions, where solutions of chaotropic salts or complexing agents like EDTA might be more promising instead and may better simulate an environmentally relevant remobilisation potential.;.
- deal with non-extractable residues in biodegradation studies: under which circumstances could NER be disregarded in the mass balance (i.e. no longer compound, but part of the matrix in practical terms), and when should they be considered part of the non-degradable fraction, because a potential for remobilisation cannot be excluded.

5.6.4 Substances of high volatility

Volatile substances can be tested according to two methods: using the closed bottle test (OECD 301 D) – however, no limit with regard to the Henry coefficient is given within OECD 301 D - or according to OECD 310 (CO_2-Headspace test) up to a Henry coefficient of 50 Pa m^3 mol^{-1}. To our knowledge, there are currently no readily applicable test systems allowing for testing of compounds of pronouncedly higher volatility

6 Conclusions

6.1 Suitability of available or proposed tests for P assessments

6.1.1 Screening tests for ready biodegradability

These tests are well established and will continue to be used as Tier 1 tests in the assessment of persistence. Several short-comings in test designs were identified, which are partly attributable to the history of test development:

1. Inoculum content is limited by the validity criteria for the blank values, which are not systematically defined.
2. A better characterisation of the inoculum with regard to the MLR of the source STP or the colony forming units could help to define the necessary amount of inoculum and lead to a higher reproducibility of tests.
3. Various attempts for pre-treatment or pre-incubation of the inoculum in order to reduce background respiration have shown to reduce the microbial diversity and potency of the inoculum.
4. Currently there are no promising routes to further standardise the inoculum (e.g. in terms of defined bacterial strains etc.)
5. The prescribed number of replicates for ready tests is low and no validity criterion exists for the variability between blank vessels. Higher replicate numbers and a validity criterion for inoculum blanks would increase reliability even in case of slightly enhanced background activity
6. Adsorption to activated sludge in DOC based tests such as OECD 301A needs careful attention; a validity criterion for elimination by adsorption should be established.
7. For investigating poorly water soluble substances reference substances (substances of poor water solubility and known biodegradability) should be established; this would improve the reliability and comparability of test outcomes for these difficult group of substances. Also, the applicability domain of tests with respect to volatile substances needs to be clearly defined.

These improvements are expected to increase the reproducibility, reliability and comparability (between different tests types, but also between laboratories carrying out the same test), and, by increasing the sensitivity (i.e. reducing the percentage of falsely not identified as biodegradable), improve the applicability of the ready tests in the assessment of persistence.

6.1.2 Enhanced screening tests for ready biodegradability

Many modifications of ready biodegradability tests are currently discussed, with the aim to improve their applicability for P assessments. Based on the discussions above we conclude:

8. Enhanced tests are screening tests and their design and interpretation should follow this categorisation: they should be of conservative nature, but should not be required to simulate environmental conditions.
9. Therefore, there is no sound argument to exclude activated sludge as an inoculum for enhanced tests (as used in standard ready tests), as suggested by the ECHA guidance (ECHA, 2014b).
10. Adaptation of the inoculum by pre-exposure to the test item is not acceptable within the definition of a screening test.
11. Co-metabolism is a relevant mechanism in real environments, but currently no promising experimental designs to include it in biodegradability screening tests are discernible due to the uncertainty when introducing an additional carbon source.
12. Test design, especially with regard to inoculum parameters (concentration, characterisation, etc.) needs standardisation and agreement (as for ready test). It is suggested that the maximum inoculum concentration should be clearly below that used for inherent tests. Instead, the total amount of the inoculum may be increased by using larger test vessels.

13. As for ready-type screening tests as discussed above, no suitable procedures for pre-treatment/pre-incubation and standardisation of inoculum are available.
14. Test duration should be limited to 60 days for all types of tests.
15. A set of reference substances should be established for enhanced ready tests, which cover various degrees of biodegradability.
16. Based on results with these reference substances in test with standardised designs agreement on suitable pass levels and validity criteria should be achieved.
17. Enhanced ready screening test should only be used for the persistency assessment (yes/no conclusion) and not for ready biodegradability evaluation. The derivation of kinetic data (half-life) is not attributable to environmental compartments.

6.1.3 Screening tests for inherent biodegradability

Some possibilities to refine or extend inherent biodegradability tests were identified, e.g.

18. Combination of the endpoints CO_2 evolution and DOC-elimination in one test to allow distinguishing between mineralisation and adsorption.
19. Performance of the MITI II test (OECD 302C) with activated sludge or other inocula to improve its applicability.
20. Addressing adsorption to activated sludge in DOC based tests (OECD 302B) by reporting the three hours value.

Despite these potential improvements it remains clear that the inherent tests are not an essential part of a generic strategy to assess persistence, but only may add relevant information in specific cases (e.g. a degradation below 20% in an inherent test may be used as indicator for a substance as persistent).

6.1.4 Compartment-specific screening tests

Newly developed compartment-specific tests (for soil and sediment) are considered screening tests by their nature. They may be promising tools for obtaining compartment-specific information on degradation, in cases where there are indications that a substance is mainly distributed to that specific compartment or that degradation in that compartment is different than in others (e.g. pelagic compartment).

In order to achieve this, further development is necessary:

21. Test designs needs standardisation and agreement (e.g. by performing a ring test).
22. It needs to be defined which is the most relevant test read-out (level of degradation, half-lives).
23. Criteria need to be established to decide on biodegradable/non-biodegradable (pass levels for degradation or half-lives in time units).
24. Their role within a testing strategy for persistence assessment should be established.

6.1.5 Simulation tests

Simulation tests are currently the last resort for biodegradability testing, because they are time- and resource intensive. Although they haven't been in the focus of this review the following issues were identified.

25. The OECD 307 test is well established for testing biodegradability in the soil department. As for any soil test the identification and evaluation of non-extractable residues is an issue that is not yet finally resolved.
26. OECD 308 and 309 would benefit from a specification of test designs and test performance criteria.

27. Clear rules should be established how to compare the total system half-life often obtained from OECD 308 with Annex XIII criteria for individual compartments.

6.2 Needs for further research and developments

Following research and development needs were identified:

- Refinement and improved standardisation of tests for (enhanced) ready biodegradability, especially with regard to inoculum parameters and validity criteria.
- Establishment of sets of reference substances (for poorly water soluble substances, for comparing results from enhanced tests and from compartment-specific tests) is important to reliably assess newly developed or amended tests.
- Improvement of screening test performance for inhibitory or low water soluble substances.
- Standardization and ring tests for compartment specific screening tests as well as for inherent type tests with CO_2 evolution.
- Critical review of the currently applied pass levels for P with regard to inherent tests (302B and C; especially when combined with CO_2 evolution) (ECHA, 2014c) which were actually developed as a prerequisite to derive half-lives for STP and environmental compartments.
- As detailed above for the particular test types, pass levels for enhanced screening tests need confirmation; and a prediction model with regard to P / not-P needs to be established for compartment specific screening tests.
- Validation of the most promising developments with enhanced screening tests by applying them to a set of reference compounds and comparing them to test results from tests for ready biodegradability and simulation tests.

7 References

Abalos, A.; Vinas, M.; Sabaté, J.; Manresa, M.A.; Solanas, A.M. (2004)

Enhanced biodegradation of Casablanca crude oil by a microbial consortium in presence of a rhamnolipid produced by *Pseudomonas aeruginosa* AT10

Biodegradation, 15, 249-260

Ahtiainen, J.; Aalto, M.; Pessala, P. (2003)

Biodegradation of chemicals in a standardized test and in environmental conditions

Chemosphere, 51, 529-537

Aronson, D.; Boethling, R.; Howard, P.; Stiteler, W. (2006)

Estimating biodegradation half-lives for use in chemical screening

Chemosphere, 63, 1953-1960

Aronstein, B.N.; Alexander, M. (1992)

Surfactants at low concentrations stimulate biodegradation of sorbed hydrocarbons in samples of aquifer sands and soil slurries

Environmental Toxicology and Chemistry, 11, 1227-1233

Baginska, E.; Haiß, A.; Kümmerer, K. (2015)

Biodegradation screening of chemicals in an artificial matrix simulating the water-sediment interface

Chemosphere, 119, 1240-1246

Barriuso, E.; Benoit, P.; Dubus, I.G. (2008)

Formation of pesticide nonextractable (bound) residues in soil: magnitude, controlling factors and reversibility

Environmental Science & Technology, 42, 1845-1854

Bartram, A.K.; Lynch, M.D.; Stearns, J.C.; Moreno-Hagelsieb, G.; Neufeld, J.D. (2011)

Generation of multimillion-sequence 16S rRNA gene libraries from complex microbial communities by assembling paired-end illumina reads

Applied and Environmental Microbiology, 77, 3846-3852

Batt, A.L.; Kim, S.; Aga, D.S. (2006)

Enhanced biodegradation of iopromide and trimethoprim in nitrifying activated sludge

Environmental Science and Technology, 40, 7367-7373

Battersby, N.S.; Ciccognani, D.; Evans, M.R.; King, D.; Painter, H.A.; Peterson, D.R.; Starkey, M. (1999)

An 'inherent' biodegradability test for oil products: description and results of an international ring test. CONCAWE Biodegradation Task Force

Chemosphere, 38, 3219-3235

Baumann, U.; Müller, M. (1996)

Ein Ansatz für kläranlagenverträgliche Produkte

Gas Wasser Abwasser, 295-299

Beek, B. (2000)

Biodegradation and persistence

Springer Verl.

Berkner, S.; Thierbach, C. (2014)

Biodegradability and transformation of human pharmaceutical active ingredients in environmentally relevant test systems

Environmental Science and Pollution Research International, 21, 9461-9467

Blok, J.; Booy, M. (1984)

Biodegradability test results related to quality and quantity of the inoculum

Ecotoxicology and Environmental Safety, 8, 410-422

Boethling, R.S.; Lynch, D.G. (2007)

Biodegradation of US premanufacture notice chemicals in OECD tests

Chemosphere, 66, 715-722

Bowmer, T.; Leopold, A. (2004)

Strategies for selecting biodegradation simulation tests and their interpretation in persistence evaluation and risk assessment. Simulation Testing of Environmental Persistence (STEP): report of a two-day workshop held in Rotterdam, 4 and 5 October, 2004

Organized by TNO - The Netherlands Organization for Applied Scientific Research, Zeist, The Netherlands; Wildlife International Ltd., Maryland, USA; Environ Corporation, Zeist, The Netherlands.
http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.486.6202&rep=rep1&type=pdf

Chang, J.S.; Radosevich, M.; Jin, Y.; Cha, D.K. (2004)

Enhancement of phenanthrene solubilization and biodegradation by trehalose lipid biosurfactants

Environmental Toxicology and Chemistry, 23, 2816-2822

Comber, M.; Holt, M. (2010)

Developing a set of reference chemicals for use in biodegradability tests for assessing the persistency of chemicals. Report No: MCC/007

http://cefic-lri.org/wp-content/uploads/uploads/Project%20publications/MCC_007_Eco12_Final_Report.pdf

CONCAWE (1999)

A test method to assess the 'inherent' biodegradability of oil products. Report no. 99/59

Prepared for CONCAWE's Petroleum Products Management Group, Brussels

Concawe (2012)

Primary biodegradation of petroleum hydrocarbons in seawater. Report no. 10/12

The oil companies' European association for Environment, Health and Safety in refining and distribution, Brussels

Cornelissen, G.; van Noort, P.C.; Govers, H.A. (1998)

Mechanism of slow desorption of organic compounds from sediments: a study using model sorbents

Environmental Science & Technology, 32, 3124-3131

Cowan, C.E.; Larson, R.J.; Feijtel, T.C.J.; Rapaport, R.A. (1993)

An improved model for predicting the fate of consumer product chemicals in wastewater treatment plants

Water Research, 27, 561-573

Cregut, M.; Jouanneau, S.; Brillet, F.; Durand, M.J.; Sweetlove, C.; Chenèble, J.C.; L'Haridon, J.; Thouand, G. (2014)

High throughput and miniaturised systems for biodegradability assessments

Environmental Science and Pollution Research, 21, 9545-9552

Cripe, C.; Walker, W.; Pritchard, P.; Bourquin, A. (1987)

A shake-flask test for estimation of biodegradability of toxic organic substances in the aquatic environment

Ecotoxicology and Environmental Safety, 14, 239-251

Czechowska, K.; Sentchilo, V.; Beggah, S.; Rey, S.; Seyfried, M.; van der Meer, J.R. (2013)

Examining chemical compound biodegradation at low concentrations through bacterial cell proliferation

Environmental Science & Technology, 47, 1913-1921

Dailey, R.; Daniel, M.; Leber, A.P. (2013)

Biodegradability of the antioxidant diaryl-p-phenylene diamine using a modified inherent biodegradation method at an environmentally relevant concentration

Chemosphere, 93, 1023-1028

Dalton, H.; Stirling, D.; Quayle, J. (1982)

Co-metabolism [and discussion]

Philosophical Transactions of the Royal Society B: Biological Sciences, 297, 481-496

EC, European Commission (1996)

Technical Guidance Document in Support of Commission Directive 93/67/EEC on Risk Assessment for New Notified Substances and Commission Regulation (EC) No 1488/94 on Risk Assessment for Existing Substances

ECSC-EC-EAEC, Brussels - Luxembourg

EC, European Commission (2003)

Technical Guidance Document in Support of the Commission Directive 93/67/EEC on Risk Assessment for New Notified Substances and the Commission Regulation (EC) 1488/94 on Risk Assessment for Existing Substances and Directive 98/8/EC of the European Parliament and of the Council Concerning the Placing of Biocidal Products on the Market

Joint Research Centre, Institute for Health and Consumer Protection, European Chemicals Bureau, Ispra, Italy

ECETOC, European Centre for Ecotoxicology and Toxicology of Chemicals (2003)

Technical Report No. 90. Persistence of Chemicals in the Environment

Brussels Belgium

ECETOC, European Centre for Ecotoxicology and Toxicology of Chemicals (2007)

Workshop Report No. 10. Workshop on Biodegradation and Persistence, 26-27 June 2007, Holmes Chapel

Brussels, Belgium

ECETOC, European Centre for Ecotoxicology and Toxicology of Chemicals (2010)

Workshop Report No. 17. Significance of bound residues in environmental risk assessment

Brussels, Belgium

ECETOC, European Centre for Ecotoxicology and Toxicology of Chemicals (2011)

Technical Report No. 112. Refined Approaches for Risk Assessment of PBT/vPvB Chemicals

Brussels, Belgium

ECETOC, European Centre for Ecotoxicology and Toxicology of Chemicals (2013a)

Technical Report No. 117. Understanding the relationship between extraction technique and bioavailability

Brussels, Belgium

ECETOC, European Centre for Ecotoxicology and Toxicology of Chemicals (2013b)

Technical Report No. 118. Development of interim guidance for the inclusion of non-extractable residues (NER) in the risk assessment of chemicals

Brussels, Belgium

ECETOC, European Centre for Ecotoxicology and Toxicology of Chemicals (2013c)

Workshop Report No. 24. Assessing Environmental Persistence, 6-7 November 2012, Paris

Brussels, Belgium

ECETOC, European Centre for Ecotoxicology and Toxicology of Chemicals (2014)

Special Report No. 18. Information to be considered in a weight-of-evidence-based PBT/vPvB assessment of chemicals (Annex XIII of REACH)

Brussels, Belgium

ECHA, European Chemicals Agency (2012)

Guidance on information requirements and chemical safety assessment. Chapter R.16: Environmental Exposure Estimation

http://echa.europa.eu/

ECHA, European Chemicals Agency (2014a)

Guidance on information requirements and chemical safety assessment. Chapter R.7a: Endpoint specific guidance. Version 3.0, August 2014

http://guidance.echa.europa.eu/

ECHA, European Chemicals Agency (2014b)

Guidance on information requirements and chemical safety assessment. Chapter R.7b: Endpoint specific guidance. Version 2.0, November 2014

http://guidance.echa.europa.eu/

ECHA, European Chemicals Agency (2014c)

Guidance on information requirements and chemical safety assessment. Chapter R.11: PBT/vPvB Assessment. November 2014 (Version 2.0)

http://echa.europa.eu/support/guidance

Ericson, J.F. (2007)

An evaluation of the OECD 308 water/sediment systems for investigating the biodegradation of pharmaceuticals

Environmental Science & Technology, 41, 5803-5811

Ericson, J.F. (2010)

Evaluation of the OECD 314B activated sludge die-away test for assessing the biodegradation of pharmaceuticals

Environmental Science & Technology, 44, 375-381

Ericson, J.F.; Smith, R.M.; Roberts, G.; Hannah, B.; Hoeger, B.; Ryan, J. (2014)

Experiences with the OECD 308 transformation test: a human pharmaceutical perspective

Integrated Environmental Assessment and Management, 10, 114-124

Federle, T.W.; Gasior, S.D.; Nuck, B.A. (1997)

Extrapolating mineralization rates from the ready Co2 screening test to activated sludge, river water, and soil
Environmental Toxicology and Chemistry, 16, 127-134

Flenner, C.K.; Parsons, J.R.; Schrap, S.M.; Opperhuizen, A. (1991)
Influence of suspended sediment on the biodegradation of alkyl esters of *p*-aminobenzoic acid
Bulletin of Environmental Contamination and Toxicology, 47, 555-560

Foladori, P.; Laura, B.; Gianni, A.; Giuliano, Z. (2007)
Effects of sonication on bacteria viability in wastewater treatment plants evaluated by flow cytometry--fecal indicators, wastewater and activated sludge
Water Research, 41, 235-243

Franklin, R.B.; Garland, J.L.; Bolster, C.H.; Mills, A.L. (2001)
Impact of dilution on microbial community structure and functional potential: comparison of numerical simulations and batch culture experiments
Applied and Environmental Microbiology, 67, 702-712

Gan, Z.; Sun, H.; Wang, R.; Hu, H.; Zhang, P.; Ren, X. (2014)
Transformation of acesulfame in water under natural sunlight: joint effect of photolysis and biodegradation
Water Research, 64, 113-122

Gao, Q.T.; Wong, Y.S.; Tam, N.F. (2011)
Removal and biodegradation of nonylphenol by immobilized *Chlorella vulgaris*
Bioresource Technology, 102, 10230-10238

Garcia-Junco, M.; Gomez-Lahoz, C.; Niqui-Arroyo, J.L.; Ortega-Calvo, J.J. (2003)
Biosurfactant- and biodegradation-enhanced partitioning of polycyclic aromatic hydrocarbons from nonaqueous-phase liquids
Environmental Science & Technology, 37, 2988-2996

Gartiser, S. (1999)
Abschätzung des Beitrages kommunaler Kläranlagen an Schadstoffkonzentrationen in Oberflächengewässern
Umweltwissenschaften und Schadstoff-Forschung, 11, 157-162

Gartiser, S.; Urich, E.; Alexy, R.; Kümmerer, K. (2007)
Ultimate biodegradation and elimination of antibiotics in inherent tests
Chemosphere, 67, 604-613

Gejlsbjerg, B.; Madsen, T.; Andersen, T.T. (2003)
Comparison of biodegradation of surfactants in soils and sludge-soil mixtures by use of 14C-labelled compounds and automated respirometry
Chemosphere, 50, 321-331

Goodhead, A.K.; Head, I.M.; Snape, J.R.; Davenport, R.J. (2014)
Standard inocula preparations reduce the bacterial diversity and reliability of regulatory biodegradation tests
Environmental Science and Pollution Research International, 21, 9511-9521

Haftka, J.J.; Parsons, J.R.; Govers, H.A.; Ortega-Calvo, J.J. (2008)

Enhanced kinetics of solid-phase microextraction and biodegradation of polycyclic aromatic hydrocarbons in the presence of dissolved organic matter
Environmental Toxicology and Chemistry, 27, 1526-1532

Hand, L.H.; Moreland, H.J. (2014)
Surface water mineralization of isopyrazam according to OECD 309: observations on implementation of the new data requirement within agrochemical regulation
Environmental Toxicology and Chemistry, 33, 516-524

Handley, J.W.; Mead, C.; Rausina, G.A.; Waid, L.J.; Gee, J.C.; Herron, S.J. (2002)
The use of inert carriers in regulatory biodegradation tests of low density poorly water-soluble substances
Chemosphere, 48, 529-534

Honti, M.; Fenner, K. (2015)
Deriving persistence indicators from regulatory water-sediment studies - opportunities and limitations in OECD 308 data
Environmental Science & Technology
49, 5879-5886

Horvath, R.S. (1972)
Microbial co-metabolism and the degradation of organic compounds in nature
Bacteriological Reviews, 36, 146-155

Howard, P.H. (1993)
Progress report: determining "real world" biodegradation rates
Environmental Toxicology and Chemistry, 12, 1135-1137

Ingerslev, F.; Nyholm, N. (2000)
Shake-flask test for determination of biodegradation rates of ^{14}C-labeled chemicals at low concentrations in surface water systems
Ecotoxicology and Environmental Safety, 45, 274-283

Ingerslev, F.; Torång, L.; Loke, M.L.; Halling-Sorensen, B.; Nyholm, N. (2001)
Primary biodegradation of veterinary antibiotics in aerobic and anaerobic surface water simulation systems
Chemosphere, 44, 865-872

Ingerslev, F.; Torång, L.; Nyholm, N. (2000)
Importance of the test volume on the lag phase in biodegradation studies
Environmental Toxicology and Chemistry, 19, 2443-2447

Jablonowski, N.D.; Köppchen, S.; Hofmann, D.; Schäffer, A.; Burauel, P. (2009)
Persistence of ^{14}C-labeled atrazine and its residues in a field lysimeter soil after 22 years
Environmental Pollution, 157, 2126-2131

Jaworska, J.S.; Boethling, R.S.; Howard, P.H. (2003)
Recent developments in broadly applicable structure-biodegradability relationships
Environmental Toxicology and Chemistry, 22, 1710-1723

Jensen, K.; Albrechtsen, H.-J.; Nielsen, J.; Kruse, B. (1988)

The use of ecocores to evaluate biodegradation in marine sediments
Water, Air, and Soil Pollution, 39, 89-99

Jiang, Z.; Yang, H.; Sun, L.; Shi, S. (2002)
Integrated assessment for aerobic biodegradability of organic substances
Chemosphere, 48, 133-138

Junker, T.; Coors, A.; Schüürmann, G. (2016)
Development and application of screening tools for biodegradation in water-sediment systems and soil
The Science of the Total Environment, 544, 1020-1030

Junker, T.; Paatzsch, C.; Knacker, T. (2010)
A water-sediment screening tool for measuring biodegradation of organic chemicals
The Science of the Total Environment, 408, 3803-3810

Kaiser, K.L.E. (1998)
Review of biodegradability tests for the purpose of developing regulations
Water Quality Research Journal of Canada, 33, 185-211

Kalsch, W.; Knacker, T.; Danneberg, G.; Studinger, G.; Franke, C. (1999)
Biodegradation of [^{14}C]-4-nitrophenol in a sediment-water simulation test
International Biodeterioration & Biodegradation, 44, 65-74

Kappesser, S.; Kutzner, H.-J. (1991)
Belebtschlamm-Agar: Obligatorisch für die Kultivierung einiger autochtoner Bakterien.
Gas Wasser Abwasser, 132, 256-257

Kästner, M.; Nowak, K.M.; Miltner, A.; Trapp, S.; Schaeffer, A. (2014)
Classification and modelling of nonextractable residue (NER) formation of xenobiotics in soil-a synthesis
Critical Reviews in Environmental Science and Technology, 44, 2107-2171

Kayashima, T.; Taruki, M.; Katagiri, K.; Nabeoka, R.; Yoshida, T.; Tsuji, T. (2014)
Comparison of biodegradation performance of OECD test guideline 301C with that of other ready biodegradability tests
Environmental Toxicology and Chemistry, 33, 328-333

Khaleel, N.D.; Mahmoud, W.M.; Hadad, G.M.; Abdel-Salam, R.A.; Kümmerer, K. (2013)
Photolysis of sulfamethoxypyridazine in various aqueous media: aerobic biodegradation and identification of photoproducts by LC-UV-MS/MS
Journal of Hazardous materials, 244-245, 654-661

Knightes, C.D.; Peters, C.A. (2006)
Multisubstrate biodegradation kinetics for binary and complex mixtures of polycyclic aromatic hydrocarbons
Environmental Toxicology and Chemistry, 25, 1746-1756

Kowalczyk, A.; Martin, T.J.; Price, O.R.; Snape, J.R.; van Egmond, R.A.; Finnegan, C.J.; Schäfer, H.; Davenport, R.J.; Bending, G.D. (2015)
Refinement of biodegradation tests methodologies and the proposed utility of new microbial ecology techniques

Ecotoxicology and Environmental Safety, 111, 9-22

Lin, A.Y.; Lin, C.A.; Tung, H.H.; Chary, N.S. (2010)

Potential for biodegradation and sorption of acetaminophen, caffeine, propranolol and acebutolol in lab-scale aqueous environments

Journal of Hazardous materials, 183, 242-250

Martin, T.J. (2014)

The influence of microbial inocula on biodegradation outcome towards enhanced regulatory assessments. Thesis submitted to Newcastle University. School of Civil Engineering and Geosciences Newcastle University, Newcastle upon Tyne, NE1 7RU, Submission: February 2014

Meinecke, S.; Norr, C.; Brackemann, H. (2000)

Prüfung der biologischen Abbaubarkeit von schwerlöslichen und oberflächenaktiven Stoffen in einem neuen Testsystem auf Basis des Zahn-Wellens-Tests.

Vom Wasser, 95, 177-190

Menon, R. (2014)

Repeat of Previously Conducted Ready Biodegradability Studies Following Current Regulatory Guidance, and its Effect on Persistence Assessment

presented at SETAC Europe 24th Annual Meeting, Basel, Switzerland, 11-15 May 2014

Mezzanotte, V.; Bertani, R.; Degli Innocenti, F.; Tosin, M. (2005)

Influence of inocula on the results of biodegradation tests

Polymer Degradation and Stability, 87, 51-56

Moltmann, J.F.; Gartiser, S. (2001)

Vergleich der Ergebnisse zum biologischen Abbauverhalten von Chemikalien aus Screening-Tests und Simulationstests für Oberflächengewässer - Teilvorhaben 2. FuE-Vorhaben FKZ 298 67 404

ECT Oekotoxikologie GmbH, Flörsheim am Main und Hydrotox GmbH, Freiburg. Im Auftrag des Umweltbundesamtes, Berlin

Mukred, A.M.; Hamid, A.A.; Hamzah, A.; Yusoff, W.M. (2008)

Enhancement of biodegradation of crude petroleum-oil in contaminated water by the addition of nitrogen sources

Pakistan Journal of Biological Sciences, 11, 2122-2127

OECD, Organisation for Economic Co-Operation and Development (1995)

Detailed Review Paper On Biodegradability Testing. OECD Series on the Test Guidelines Programme Number 2. Environment Monograph N° 98

Paris, France. http://www.oecd.org/chemicalsafety/testing/seriesontestingandassessmentpublicationsbynumber.htm

OECD, Organisation for Economic Co-Operation and Development (2016)

OECD Guidelines for the Testing of Chemicals

http://www.oecd-ilibrary.org/environment/oecd-guidelines-for-the-testing-of-chemicals_chem_guide_pkg-en

OECD/OCDE (2006)

OECD Guidelines for the Testing of Chemicals. Revised Introduction to the OECD Guidelines for Testing of Chemicals, Section 3. Part 1: Principles and Strategies Related to the Testing of Degradation of Organic Chemicals. Adopted: 23 March 2006

http://www.oecd-ilibrary.org/docserver/download/9730001e.pdf?expires=1358164047&id=id&accname=guest&checksum=13C836D81324F2E6E32FF255AA0709F4

Paixão, S.M.; Sàágua, M.C.; Tenreiro, R.; Anselmo, A.M. (2006)
Biodegradability testing using standardized microbial communities as inoculum
Environmental Toxicology, 21, 131-140

Parsons, J.R. (1992)
Influence of suspended sediment on the biodegradation of chlorinated dibenzo-p-dioxins
Chemosphere, 25, 1973-1980

Radke, M.; Maier, M.P. (2014)
Lessons learned from water/sediment-testing of pharmaceuticals
Water Research, 55, 63-73

Rauert, C.; Friesen, A.; Hermann, G.; Jöhncke, U.; Kehrer, A.; Neumann, M.; Prutz, I.; Schönfeld, J.; Wiemann, A.; Willhaus, K.; Wöltjen, J.; Duquesne, S. (2014)
Proposal for a harmonised PBT identification across different regulatory frameworks
Environmental Sciences Europe, 26:9, 1-12

Richterich, K.; Steber, J. (2001)
The time-window - an inadequate criterion for the ready biodegradability assessment of technical surfactants
Chemosphere, 44, 1649-1654

Scheringer, M.; Macleod, M.; Matthies, M.; Klasmeier, J.; Witt, J. (2006)
Persistence Criteria in the REACH Legislation: Critical Evaluation and Recommendations. Final Report, 21.12.2006
http://www.reach-info.de/dokumente/gutachten_gesamtpersistenz.pdf

Schlötelburg, C. (2001)
Mikrobielle Diversität und Dynamik einer 1,2-Dichlorpropan dechlorierenden Mischkultur. Dissertation zur Erlangung des akademischen Grades doctor rerum naturalium im Fach Biologie eingereicht an der Mathematisch-Naturwissenschaftlichen Fakultät I der Humboldt-Universität zu Berlin

Seyfried, M.; van Ginkel, C.G.; Boschung, A.; Miffon, F.; Fantini, P.; Tissot, E.; Baroux, L.; Merle, P.; Chaintreau, A. (2015)
Persistence assessment of cyclohexyl- and norbornyl-derived ketones and their degradation products in different OECD screening tests
Chemosphere, 131, 63-70

She, Z.; Xie, T.; Zhu, Y.; Li, L.; Tang, G.; Huang, J. (2012)
Study on the aerobic biodegradability and degradation kinetics of 3-NP; 2,4-DNP and 2,6-DNP
Journal of Hazardous materials, 241-242, 478-485

Shimp, R.J.; Schwab, B.S.; Larson, R.J. (1989)
Adaptation to a quaternary ammonium surfactant by suspended microbial communities in a model stream
Environmental Toxicology and Chemistry, 8, 723-730

Shin, K.H.; Ahn, Y.; Kim, K.W. (2005)
Toxic effect of biosurfactant addition on the biodegradation of phenanthrene
Environmental Toxicology and Chemistry, 24, 2768-2774

Smith, K.E.; Rein, A.; Trapp, S.; Mayer, P.; Karlson, U.G. (2012)
Dynamic passive dosing for studying the biotransformation of hydrophobic organic chemicals: microbial degradation as an example
Environmental Science & Technology, 46, 4852-4860

Snaidr, J.; Amann, R.; Huber, I.; Ludwig, W.; Schleifer, K.H. (1997)
Phylogenetic analysis and in situ identification of bacteria in activated sludge
Applied and Environmental Microbiology, 63, 2884-2896

Solomon, K.; Matthies, M.; Vighi, M. (2013)
Assessment of PBTs in the European Union: a critical assessment of the proposed evaluation scheme with reference to plant protection products
Environmental Sciences Europe, 25, 10

Spain, J.C.; Van Veld, P.A. (1983)
Adaptation of natural microbial communities to degradation of xenobiotic compounds: effects of concentration, exposure time, inoculum, and chemical structure
Applied and Environmental Microbiology, 45, 428-435

Sparham, C.; Rehman, N.; Melling, J.; van Duynhoven, J.; Marshall, S. (2008)
Biodegradability of highly ethoxylated nonionic surfactants: determination of intermediates and pathways of biodegradation
Environmental Toxicology and Chemistry, 27, 1069-1076

Stasinakis, A.S.; Petalas, A.V.; Mamais, D.; Thomaidis, N.S. (2008)
Application of the OECD 301F respirometric test for the biodegradability assessment of various potential endocrine disrupting chemicals
Bioresource Technology, 99, 3458-3467

Strotmann, U.J.; Schwarz, H.; Pagga, U. (1995)
The combined CO_2/DOC test - a new method to determine the biodegradability of organic compounds
Chemosphere, 30, 525-538

Strotmann, U.J.; Weberruß, U.; Bias, W.R. (1993)
Degradation of morpholine in several biodegradation tests and in wastewater treatment plants
Chemosphere, 26, 1729-1742

Struijs, J.; Stoltenkamp-Wouterse, M.J.; Dekkers, A.L. (1995)
A rationale for the appropriate amount of inoculum in ready biodegradability tests
Biodegradation, 6, 319-327

Thouand, G.; Capdeville, B.; Block, J.C. (1996)
Preadapted inocula for limiting the risk of errors in Biodegradability Tests
Ecotoxicology and Environmental Safety, 33, 261-267

Thouand, G.; Durand, M.J.; Maul, A.; Gancet, C.; Blok, H. (2011)
New concepts in the evaluation of biodegradation/persistence of chemical substances using a microbial inoculum
Frontiers in Microbiology, 2, 164

Thouand, G.; Friant, P.; Bois, F.; Cartier, A.; Maul, A.; Block, J.C. (1995)
Bacterial inoculum density and probability of *para*-nitrophenol biodegradability test response
Ecotoxicology and Environmental Safety, 30, 274-282

Toräng, L.; Nyholm, N.; Albrechtsen, H.J. (2003)
Shifts in biodegradation kinetics of the herbicides MCPP and 2,4-D at low concentrations in aerobic aquifer materials
Environmental Science & Technology, 37, 3095-3103

Torres-Bojorges, A.X.; Buitrón, G. (2012)
Biodegradation of nonylphenols using nitrifying sludge, 4-chlorophenol-adapted consortia and activated sludge in liquid and solid phases
Environmental Technology, 33, 1727-1737

UBA, Umweltbundesamt (2015)
Erarbeitung eines gestuften Extraktionsverfahrens zur Bewertung gebundener Rückstände
Mitteilung des Ausschusses für Innenraumluftwerte

Van der Harst, J.; Schmidt-Bleek, F.; De Kreuk, J.; Günther, K. (1981)
Conclusions and recommendations resulting from the secotox workshop "Feasibility of Integrated Approaches to Testing of Chemicals, Kyoto, April 21-23, 1980"
Ecotoxicology and Environmental Safety, 5, 377-381

van Ginkel, C.G.; Gancet, C.; Hirschen, M.; Galobardes, M.; Lemaire, P.; Rosenblom, J. (2008)
Improving ready biodegradability testing of fatty amine derivatives
Chemosphere, 73, 506-510

Vazquez-Rodriguez, G.; Goma, G.; Rols, J.L. (2000)
Toward a standardization of the microbial inoculum for ready biodegradability testing of chemicals
Water Science and Technology, 42, 43-46

Vazquez-Rodriguez, G.; Goma, G.; Rols, J.L. (2003)
Activated sludge as inoculum for ready biodegradability testing: effect of source
Environmental Technology, 24, 979-987

Vázquez-Rodríguez, G.; Palluy, F.; Goma, G.; Rols, J.L. (1999)
Procedures in ready biodegradability testing: effects of the inoculation and the monitored parameter
Environmental Technology, 20, 301-308

Vázquez-Rodríguez, G.A.; Beltrán-Hernández, R.; Coronel-Olivares, C.; Rols, J.L. (2011)
Standardization of activated sludge for biodegradation tests
Analytical and Bioanalytical Chemistry, 401, 1127-1137

Vázquez-Rodríguez, G.A.; Calmon, A.; Silvestre, F.; Goma, G.; Rols, J.L. (2006)
Effect of the inoculation level in aerobic biodegradability tests of polymeric materials
International Biodeterioration & Biodegradation, 58, 44-47

Vázquez-Rodríguez, G.A.; Garabétian, F.; Rols, J.L. (2007)

Inocula from activated sludge for ready biodegradability testing: Homogenization by preconditioning
Chemosphere, 68, 1447-1454

Watson, H.M. (1993)
A comparison of the effects of two methods of acclimation on aerobic biodegradability
Environmental Toxicology and Chemistry, 12, 2023-2030

Wenzel, W. (2002)
Mikrobiologische Charakterisierung eines Anaerobreaktors zur Behandlung von Rübenmelasseschlempe. Dissertation zur Erlangung des akademischen Grades Doktor der Naturwissenschaften
Technische Universität Berlin

Wiesmann, U.; Choi, I.S.; Dombrowski, E.-M. (2007)
Fundamentals of biological wastewater treatment.
Wiley-VCH Verl. Weinheim

Wos, M.L. (2005)
Methods For Understanding Bacterial Metabolic Activity In Activated Sludge. Thesis (PhD Doctorate)
http://research-hub.griffith.edu.au/display/n319b22f694755a30ca11a84734e14ea3

Wu, W.Z.; Xu, Y.; Schramm, K.W.; Kettrup, A. (1997)
Study of sorption, biodegradation and isomerization of HCH in stimulated sediment/water system
Chemosphere, 35, 1887-1894

Xia, X.; Wang, R. (2008)
Effect of sediment particle size on polycyclic aromatic hydrocarbon biodegradation: importance of the sediment-water interface
Environmental Toxicology and Chemistry, 27, 119-125

Xia, X.; Yu, H.; Yang, Z.; Huang, G. (2006)
Biodegradation of polycyclic aromatic hydrocarbons in the natural waters of the Yellow River: Effects of high sediment content on biodegradation
Chemosphere, 65, 457-466

Xia, X.; Zhou, Z.; Zhou, C.; Jiang, G.; Liu, T. (2011)
Effects of suspended sediment on the biodegradation and mineralization of phenanthrene in river water
Journal of Environmental Quality, 40, 118-125

8 Annex: Workshop documentation: Biodegradability tests for assessing persistence – Where are we?

Projekt Nr. 54429

Workshop and Expert Discussion

Biodegradability tests for assessing persistence –

Where are we?

16th - 17th February 2016 Umweltbundesamt, Dessau-Roßlau, Germany

Workshop documentation

Stefan Gartiser
Hydrotox GmbH, Freiburg / Breisgau

Klaus Schneider
Forschungs- und Beratungsinstitut Gefahrstoffe (FoBiG) GmbH, Freiburg /Breisgau

Markus A. Schwarz
Forschungs- und Beratungsinstitut Gefahrstoffe (FoBiG) GmbH, Freiburg /Breisgau

Thomas Junker
ECT Oekotoxikologie GmbH, Flörsheim / Main

on behalf of Umweltbundesamt

February 2016

Table of contents

Introduction

Biodegradability tests and PBT/vPvB assessments

Screening tests for evaluating persistence in the context of PBT/vPvB assessments under REACH are purposely conservative. In consequence, many substances require further investigations, when having been tested negative in screening tests. Efforts are currently made to further develop ("enhance") available tests or to develop new ones, which provide reliable conclusions on persistence without the difficulties associated with performance of simulation tests.

UBA-initiated literature study on deficits of biodegradability tests and recent developments

A literature study was performed by a group of German consultancies (FoBiG, Freiburg; Hydrotox, Freiburg; ECT, Flörsheim) to review the applicability of established and new tests for biodegradability for assessing persistence. Screening tests for ready and inherent biodegradability, proposals for enhanced screening tests as well as newly designed compartment-specific screening tests were analysed for strengths and weaknesses and proposals are made how to improve their performance and suitability for assessments of persistence. Although not in the focus of this study, some recommendations are also given for simulation tests (OECD 307, 308, 309) in the context of evaluating persistence.

Workshop and expert discussion

Within a 2-(half)-days' workshop and expert discussion at the venue of the German Umweltbundesamt (UBA) participants were informed about the project results in short presentations. Each presentation was followed by a discussion. At the end of each day round table discussions were performed. Important results are documented here to be made available to the participants.

1 Background: Regulatory requirements for assessment of persistence

1.1 Presentation

Markus Schwarz (FoBiG) gave an overview on the regulatory background for assessment of persistence. He described the intentions of REACH Annex XIII, exact criteria for persistence according to Annex XIII, and finally on the possible tests with regard to persistence and the evaluation thereof. On this behalf, relevant statements of REACH Annex XIII as well as applicable ECHA guidance documents were summarized, with special emphasis on enhanced screening test as well as difficulties in interpretation of simulation tests. A further topic was the requirements for difficult substances, namely UVCBs as well as poorly soluble substances, and associated difficulties for persistence assessment under consideration of ECHA guidance documents.

1.2 Discussion

No important discussion points were identified following this presentation.

2 Screening tests for ready biodegradation: deficits and possible improvements

2.1 Presentation

Stefan Gartiser (Hydrotox GmbH) presented the main results of the study with regard to the existing standard screening tests. The inoculum concentration in these tests is a critical point, because it differs by a factor of 10,000 between the tests combined in OECD 301. The maximum inoculum level of the different tests is determined by their validity criteria for inoculum blanks, which are not consistent and should be revised. There is an option for a better description of the inoculum (e.g. concerning the mass loading rate of the STP). Any pre-treatment and pre-incubation of the inoculum results in reduced activity and should be avoided as far as possible.

Another point of discussion is the number of replicate test and blank vessels, which should be increased to a minimum of 3 replicates, in order to describe the variability of the degradation extents and the inoculum blanks, which are directly considered in the calculation. For this, validity criteria for the variability of inoculum blanks (e.g. CV ≤ 20%) should be defined.

With respect to the DOC elimination based tests (OECD 301 A and E) no validity criteria for the inoculum blank exist, but should also be defined (e.g. 10% of TOC introduced), following the example of the corresponding ISO guidelines. Further on, the criteria for abiotic DOC elimination through adsorption or volatility should be clearly defined (e.g. ≤ 20%). The obligatory consideration of abiotic controls is another option for avoiding abiotic losses, being interpreted as biodegradation.

Alternative test approaches proposed in literature are often limited in their interpretability, because they often are not in line with basic OECD tests principles (e.g. use of supernatant activated sludge used, low test vessel volumes, biomass growth as an endpoint). The relevance of these approaches for the improvement of standard tests for regulatory purposes remains unclear.

With respect to the inherent tests, for the Zahn-Wellens test (OECD 302) it is recommended to define clear criteria on how adsorption to activated sludge should be considered (e.g. maximum allowed after 3 h ≤ 20%). The applicability of the MITI(II) OECD 302C test with other inoculum than the mixed and pre-incubated inoculum from 10 sites should be reconsidered. Further on, a standard for an inherent CO_2-evolution test should be developed.

2.2 Discussion

In the discussion there was a broad agreement that the standard screening tests could be improved and should be revised. It was suggested to prioritize the different proposals for improvements. From the consultants point of view the most important revision needed concerns the validity criteria of inoculum blanks, because the pre-incubation of the inoculum, which for some tests is needed to meet the criteria, causes a significant reduction of the activity. This is also an important issue with regard to the enhanced screening tests. Other options, such as the increase of the replicate vessels, could be (and are already in some laboratories) realized today and do not need a revision of the guideline.

With respect to the inherent tests it was concluded, that these have very favourable test conditions for pre-adaptation. Thus it remains unclear, how data from the Zahn-Wellens test could be used for the “P”-assessment. For the same reason, inherent test conditions should not be considered acceptable for enhanced tests.

It was suggested, that the recommendations should be implemented in an OECD update and all methods should be more harmonized. The extensions could also be useful for the performance of enhanced screening tests. In a second step the better use of ready and enhanced screening tests for the P-assessment should be assessed.

3 Screening tests for enhanced biodegradation: current developments

3.1 Presentation

In the next presentation Stefan Gartiser presented the proposals for „enhanced screening tests". He recommended that it should clearly be stated that the enhanced screening tests refer to enhanced "ready type" tests, while "inherent type" enhanced screening tests do not exist. There are enhancements, which rather represent an improvement of the test performance, without losing their attribution as "ready type" tests. Other enhancements which are not already described in OECD 301 guidelines cannot be used for ready biodegradability testing, but exclusively for the "non-P" assessment. Among the first, the use of larger test vessels to increase the starting biomass is in line with the OECD 301 guidelines and therefore no upper limit should be prescribed. The inclusion of further measurements (DOC, chemical analytics), the characterization of the inoculum diversity and biomass growth, or the establishment of "cold" carbon balances (non-labelled carbon balance while considering CO_2-evolution, DOC in the water phase, biomass growth, and residual polymers) may be reasonable. A set of positive and negative reference compounds and the increase of replicate vessels are other options, which also for enhanced tests improve the accuracy of the tests. Attempts for a standardization or pre-treatment of the inoculum usually is associated with reduced activity and should be avoided if possible. A point for discussion was that the existing ECHA guidance, which does not allow for inoculum derived from sewage treatment plants for use in "P"-assessment (only natural environmental media are allowed). This is contradictory to the conclusions drawn from standard OECD 301 tests, which already are used to assess substances as not being "P". Further on, STP effluents as well as suspended solids from STP are continuously released into surface water. The prolongation of test duration for up to 60 days is an option already referred to in the OECD 301 guideline, but the results cannot be used for "ready biodegradable" assessments. The same is true for an increase of the inoculum concentration. In order to avoid an inherent type character of the enhanced test an upper limit of e.g. 100 mg dry solids /L activated sludge was proposed. The use of inoculum pre-adapted to the test item should not be allowed. With respect to possible combinations of enhancements it was suggested to allow all combinations which do not question the attribution as "ready type tests" and one additional enhancement (increase of inoculum concentration or prolongation beyond 28 d).

3.2 Discussion

The different proposals for enhancements were discussed:

With respect to the prolongation some participants agreed that this option is already proposed in the guideline and ECHA guidance. A prolongation could be an option for substances which show initial biodegradation, but have not reached the plateau by day 28. On the other hand, the reasons for slow biodegradation should be taken into account. If the degradation of the compound is hindered by its low water solubility (e.g. long chain hydrocarbons), there is also the option to improve the bioavailability by methods such as addition of SiO_2 gel, emulsifiers, or solvents, as already described in the OECD 301 introduction and ISO guidance. Some participants preferred such an option against a prolongation of the test. The arguments were that a slow biodegradation could also be caused by late adaptation of the inoculum to the test item. On the part of the consultants, the prolongation of a running test was preferred as a simple way for considering low bioavailability. The addition of solubility aids increase uncertainty in the tests (e.g. through inhibitory effects) and requires additional control vessels to check the effects solubility aids may have had on the test system. One participant assumed prolongation not as best enhancement, but not as critical, because of the substantial loss of the inoculum activity. In some way a prolongation to 60 days could be interpreted as roughly equal to a low

level pre-adaption suggested within ECHA guidance R.7b, namely repetition of a ready test with the (possibly adapted) sludge of a ready test run beforehand with the same test item. There was also a discussion whether the addition of solubility aids still represents environmental conditions, but the overall conclusion was that screening tests do not intend representing real environments. Some participants also agreed on combining both the addition of solubility aids and a prolongation for certain substance groups. Thus, the physico-chemical characteristics of the test item should be taken into account for a suitable testing strategy. Some participants requested a more systematic evaluation of data and validation of such enhancements.

The use of STP derived inoculum also for enhanced screening tests was agreed upon by all participants, although some concern remained, because STPs with continuous influents of all sorts of chemicals may be considered being adapted. However, the arguments for using STP inoculum are conclusive and a practical way forward. When the ECHA guidance was drafted, the main concern against STP inoculum was, that adapted inoculum from STP laboratory tests should not be used for enhanced tests.

The vessel size was not considered being a critical point, because it only elevates the volume of inoculation media in the tests, without impacting either inoculum concentration or ratio test item to inoculum.

With respect to the increase of inoculum concentration most participants were not in favour of such an option, although this is mentioned in the ECHA guidance. The main concern was that in combination with other enhancements such as the increase of the vessel size the test conditions might be too favourable for still allowing a “non-P”-assessment. One participant referred to the already existing differing inoculum concentrations in standard tests, covering a factor of 10.000. In this sense, a slight increase to 100 mg/L (factor 3) might not be critical. But most participants requested a better standardization of the inoculum concentration (e.g. similar inoculum/substrate ratios) instead of allowing higher concentrations. Any proposals for an increase of the inoculum concentration should be supported by robust data, first. Thus, there should be a decision that enhanced tests should be more related to ready type screening tests than to inherent tests.

The workshop participants unanimously agreed with not using adapted inoculum. There have been approaches where a first test with standard inoculum was performed followed by a second test with (possibly) adapted inoculum from the first (i.e. no other / new inoculum added). This resembles a test prolongation up to 60 days, however with the advantage of preventing deterioration or falling apart of single test replicates often observed over longer incubation periods. This was the reason why the option of using “preadapted sludge” has been mentioned in the ECHA guidance for enhanced tests. But also for these modifications the same concerns as for the prolongation existed among the participants. Without presenting sound data supporting this enhancement, it would not be acceptable.

A better characterization of the inoculum is an option for describing biodiversity at least retrospectively. The question arose whether the characterization of the inoculum could be integrated into screening test guidelines (at least as an option). In case of evaluating equivocal screening test results, this information possibly could be useful to explain opposing results and as a means to decide on the test(s) with higher predictability. Some authorities consider one valid positive ready test as sufficient for identifying a substance as not P; others compare the acceptability and representativeness of each test. Indeed, the different potencies of the OECD 301 tests due to the differing inoculum concentrations and pre-treatments should be considered. Guidance on when to use which ready or enhanced tests for “P” assessment is insufficient and should be improved. However, before requesting further data on the characterization of the inoculum, possible indicators for inoculum diversity / activity should be critically assessed regarding their predictive value to cost ratio.

Most participants doubt whether adding co-substrates for improving co-metabolism is a good option. The co-substrate might influence biodegradation, but the question remains whether the accuracy of test method is reduced and whether the co-substrate represents environmental conditions. Thus, adding co-substrates in screening test was not recommended. On the other hand, co-substrates are implicitly present in natural media used within simulation tests.

With respect to the evaluation of data it was discussed whether enhanced tests should have reduced or elevated pass levels compared to the ready tests. A reduced pass level was not considered being meaningful, because the presumption of readily biodegradable substances is 100% mineralization, which means that the part not detected as CO_2 evolution, oxygen consumption or DOC elimination is attributed to biomass growth. Thus, a reduction of the pass levels to e.g. 50% does not necessarily mean complete degradation, while higher pass levels might be difficult to achieve due to biomass growth. A better look on biomass growth would improve the interpretation of test data. Cold carbon balance (as e.g. in ISO 14852); suspended solids measurements; and / or biomass quantification could be used to characterise assimilation.

The question, which combinations of enhancements could be allowed, was indirectly answered by the participants. With the constraints regarding a prolongation of the test duration as well as objections against all other possible enhancements, e.g. an increase of the inoculum concentration, the only remaining option (beyond the ones within the “ready type” definition) is a prolongation under certain conditions. Other options, such as increase of vessel size and replicate vessels, additional measurements and characterization of the inoculum etc. are in line with standard ready type tests and could be performed already.

A suitable way forward would be to revise the OECD 301 test guidelines and to establish a guideline for enhanced screening tests.

4 Agreed conclusions on discussions of day 1

4.1 Ready type screening tests

There is a general agreement that ready biodegradability tests would benefit from a revision to overcome some differences and inconsistencies caused by the historical development of these tests. Improvements are considered possible with regard to the number of replicates (increase of number of test and blank replicates), consideration of blank variation (definition of a criterion for maximum CV of blanks, validity criteria for blank respiration), test parameters (harmonised inoculum criteria: total amount of inoculum and ratio of test item concentration to inoculum, definition of criterion for max. allowed adsorption for DOC-based tests) and use of reference substances.

4.2 Enhanced screening tests

Increasing test vessel size is a modification which is considered suitable without leading to a change of the character of the test.

Prolongation of test duration is considered an option under certain conditions: most of the workshop participants considered this being acceptable for substances of low bioavailability, based on consideration of physico-chemical properties (e.g. water solubility) and structure and where a steady, slow degradation without reaching a plateau is observed within 28 days. Otherwise the test duration should not be prolonged, because a late increase of degradation could be caused by adaptation of bacteria. Further definition of applicability criteria and more experimental results produced under these conditions are needed.

But for substances of low bioavailability increasing bioavailability (e.g. by silica gel matrices) is the first option, which could also be combined with test prolongation.

Requirements for inoculum concentrations and conditions in enhanced biodegradability tests should be the same as those for tests for ready biodegradability (see suggestions above for harmonisation). Use of suitable positive and negative reference substances is also recommended to test performance of enhanced tests. Use of municipal STP inoculum (without adaptation in the laboratory) is acceptable also in enhanced tests and the respective part of the ECHA Guidance R.7b should be changed.

Pre-adaption is not a modification finding regulatory acceptance.

Ready and enhanced tests would benefit from making available more information on inoculum used (mass loading rate, sludge retention time) and measurement of additional parameters such as DOC, biomass growth and/or carbon balance. Such information should be considered to be included in Guidances/Guidelines (as optional parameters).

5 Compartment-specific screening tests: recent developments

5.1 Presentation

Thomas Junker (ECT) described the development and state of the art for compartment specific screening tests with sediment and soil.

More recent developments for sediment (Baginska et al., 2015; Junker et al., 2016; Junker et al., 2010) are based on artificial sediment similar to OECD 218, MITI-I based inoculum and measurements of O_2 consumption. The test is intended to be equivalent to ready biodegradability tests but incorporating sediment. As such, results with sample compounds were compared with MITI-I test results. Based on a low coefficient of variation of 6% for the mineralization of the reference compound aniline at the plateau (n= 18), the test was concluded to be applicable.

Most recently, also a soil screening test was developed and first results are available (Junker et al., 2016). In this test, natural LUFA standard soil (sandy loam) is used, thus the inoculum is media-specific and endemic. Like in the sediment screening test, biodegradation is monitored via oxygen consumption and correspondingly, high substrate concentrations typical for screening tests are applied for a sufficient signal to noise ratio. Also for this test, coefficient of variation for mineralization at the plateau was rather low (11.8% for reference compound sodium benzoate, n= 20), such that also this test was considered to be applicable.

In addition, the experimental results from water-sediment and soil screening tests (e.g. mineralization rates, and mineralization half-lives) were in good accordance with degradation data from other tests with water-sediment or soil reported in databases and literature.

5.2 Discussion

Dominating topic of debate was on the purpose of these tests and their role in P assessment. Under which circumstances do they bring relevant additional information? Under which conditions should they be used?

While tests on ready biodegradability cannot be replaced by these tests, they may be too artificial as to substitute simulation tests on sediment and soil, mostly due to the high test concentrations generally necessary for screening type tests. It is thus not clear, how either increased or decreased biodegradation results found in media specific screening tests compared to tests on ready biodegradability should be judged.

On the other hand, media specificity of the tests would imply the need to identify the compartment(s) of concern, and critical aspects in this regard were discussed: in a scenario with dominating emission into waste water passing sewage treatment plants, compounds with PBT-like properties will largely adsorb to sludge and via the pathway sludge to agricultural fields, soil would be the compartment of highest concern. Applying multi-media modelling, mass distribution is highly dependent on the emission scenario and size of compartments, and sediment may turn out as the compartment of highest concern. Generally, knowledge on emission scenarios is too limited to allow sound conclusions. Water was suggested as generally being the most important compartment due to its "moving" character and its property to distribute chemicals to/between other compartments. At the same time, for many chemicals emission into water is most important. Further, even for chemicals prone to adsorption and distribution to sediments, sedimentation rates depend on water flow rate and depth of the water column (default assumption for risk assessment: 2 m) and may thus take a rather long time.

The sediment screening test was critically discussed because degradation observed within a test implying anaerobic mechanisms as possibly decisive degradation route may not suffice to demonstrate non-persistence; because aerobic environments are generally dominating, aerobic biodegradability was regarded as inevitable property for a non-persistent compound by some of the workshop participants. On the contrary, a compound assessed to be anaerobically non-biodegradable but very well biodegradable aerobically were not to be regarded as persistent. However, for highly adsorptive compounds, it might be worth considering anaerobic biodegradability in the future.

6 Simulation tests: current status and outlook

6.1 Presentation

Thomas Junker described existing simulation tests on soil (OECD 307), sediment (OECD 308) and water (OECD 309); their purpose; associated difficulties; as well as the outcome of current research projects (Cefic LRI-ECO18 project) and associated need for improvements and further research.

With regard to the soil simulation test (OECD 307), essentially no inadequacies of the test itself were identified; rather, evaluation of test results poses difficulties: high variation of degradation rates of a test compound between different soils and uniform Annex XIII cut-off for half-life in soil; how to define, determine and interpret non-extractable residues (NER)?

In contrast, the sediment simulation test is being criticized for several reasons:

- Inappropriate ratio water : sediment between 3 : 1 and 4 : 1, being rather characteristic for small ditches relevant for field run-off of plant protection products than for surface waters impacted by chemicals relevant under REACH;
- Static design, not representing natural conditions with regard to flow velocity and sediment dynamics;
- Two-phase system of water (oxygenated) and sediment (thin aerobic layer, bulk anaerobic layer), precluding derivation of reliable apparent first order half-lives for water and sediment, which are actually required for comparison with Annex XIII cut-offs.

To overcome these problems and for a deeper understanding of relevant factors driving degradation and phase partitioning, within the Cefic LRI-ECO18 project the standard OECD 308 set-up was compared to a modified OECD 308 with a water to sediment ration of 10:1 and stirred water phase (stratified); a modified OECD 309 (higher sediment ratio, i.e. water : sediment 100:1; stirred and mixed) as well as OECD 309 (water : sediment ratio of 1000:1, stirred and mixed). Important conclusions given are

- $DegT_{50,\ system}$ is most reliable, and generally to be preferred over DT_{50} for sediment and water. Still, it is partly system dependent.
- Increasing surface area water-sediment and increasing oxygenation does lead to both, higher degradation rates and increased formation of non-extractable residues.

With regard to OECD 309, the pelagic test is expected to be of low potency such that few substances would pass regulatory P-criteria (i.e. not P). With regard to the variant of OECD 309 including sediment, shaking should be used instead of stirring to prevent extensive grinding of sediment (sand to silt/clay) associated with increased adsorption / NER formation. A variety of experimental options possible in this test hampers comparison across compounds and tests.

6.2 Discussion

Generally, simulation tests with soil (OECD 307) and sediment (OECD 308) often pose the problem of non-extractable residues (NER). As there is no agreed solution to the problem of how to analyse and interpret NER, several participants proposed to avoid those tests and rather perform the surface water simulation test (OECD 309) instead, as far as possible. This was also suggested for e.g. pharmaceuticals, which often are adsorbing but not hydrophobic as well as for adsorbing cationic substance: due to continuous release to water and water as the one "moving phase" it would be of special importance. Further, transfer to sediment depended on sedimentation rate and water depth and thus were often rather slow.

With regard of the stated low potency of the pelagic OECD 309 compared to e.g. the water-sediment test (OECD 308), the question was raised by a participant as to what has been compared: Had non-extractable residues formed within OECD 308 been counted as "degraded" or "non-degraded"? Only in case of the latter, such a comparison would be valid.

While several experimental options possible within OECD 309 may hamper inter-test comparison, there was the opinion that this flexibility is important and should be kept to enable to tailor the test according to test item properties or relevant environmental conditions. Rather, ECHA guidance documents could outline under what circumstances which options should reasonably be chosen to the effect that similar test conditions are used under similar starting conditions/for similar compounds. It was agreed upon, that for a higher number of substances results according to OECD 309 are needed to gain sufficient experience with this test. Nonetheless, the test was considered to provide suitable results for the water compartment. Further, only one specimen of surface water tested within OECD 309 was judged insufficient: with regard to the two sediment samples required for OECD 308, at least two different surface water samples should be tested. This should preferably be fixed in the ECHA guidance documents rather than in the test guideline.

Concerning OECD 308 (water-sediment) test, participants agreed to use the apparent first order total system degradation half-life ($DegT_{50\text{-}system}$) rather than non-reliable half-lives for water and sediment. For candidate PBT compounds it were to be expected that due to general hydrophobicity partitioning into the sediment phase is rather quantitative and $DegT_{50\text{-}system}$ could be compared to the regulatory half-life for sediment.

For evaluation of soil simulation test (OECD 307) half-lives on different soils and comparison to regulatory cut-off values, participants agreed to include suggestions given by Rauert et al. (2014) into ECHA guidance documents. Further, with regard to extraction schemes for NER, a proposal found agreement to specially account for cationic adsorbing compounds which might not be extractable by organic solvents generally applied but rather using high salt solutions possibly combined with chaotropic agents.

6.3 Discussion of other, more general aspects

The importance for identification of degradation products was discussed: it was emphasized by a participant that simulation tests would be the only way to accomplish this, and these tests would also be demanded to with the objective of identifying degradation products. On the other hand, in case that a substance passed an enhanced screening test, alike other screening tests mineralization may be assumed and no identification of degradation products would be needed. It became however obvious, that more detailed guidance is required as to when and how identification of degradation products is needed. It was noted that importance of degradation products is not only for PBT but also for risk assessment when degradation products are of pronouncedly higher toxicity than the parent (e.g. ethoxylated nonylphenol and the transformation product, nonylphenol).

7 Agreed conclusions on discussions of day 2

7.1 Compartment-specific screening tests

There are reservations against use of these new tests regarding their potential role (under which circumstances can they provide additional helpful information for the P assessment? Which compartment should be tested?). It was also mentioned that in the sediment test mainly anaerobic conditions prevail, which are not driving the P assessment.

These reservations are explicitly expressed for the use for P assessment and do not consider potential uses for compartment specific risk assessment.

7.2 Simulation tests

7.2.1 OECD 307 (soil simulation test)

In the ECHA Guidance, recommendations should be given how to deal with results from various soils tested. This should be done by (beyond others) including the following suggestions based on the work of Rauert et al. (2014):

- In case of results for up to 4 different soils (comparable reliability assumed): use worst-case $DegT_{50}$-value,
- In case of results for 5 and more different soils of comparable reliability: use value between median and 90th percentile.

7.2.2 OECD 309 (surface water simulation test)

Few experimental results exist for this test, yet. However, the test is considered to provide suitable results for the water compartment. Two different surface water samples should be used as a minimum, and this should be recommended in the ECHA Guidance. Participants consider this test especially suitable for substances, which may lead to high NER (non-extractable residue) in sediment or soil tests. Reference substances should be used to improve / test its performance.

Harmonisation of test conditions: For the optional addition of sediment, shaking instead of stirring is recommended by the workshop participants to prevent grinding of sediment and associated higher adsorption / NER formation.

For very hydrophobic substances (with low water solubility) the 309 test could be performed with addition of (suspended) sediment. However, care should be taken if NER formation is expected.

7.2.3 Conclusion on 308 (water-sediment simulation test)

For P assessment mostly substances with high hydrophobicity (e.g. sediment distribution coefficient, K_d > 2000 L/kg) are relevant. For these substances the total system $DegT_{50}$ should be compared to the sediment-specific cut-off value in Annex XIII, REACH. If $DegT_{50\text{-system}}$ value is used, test vessel geometry should be reported.

7.2.4 General aspects on simulation tests

Harmonisation of test conditions (e.g. system geometry for OECD 308; for OECD 309: shaking in case sediment is included; definition of light intensity in case of light-dark cycles included; minimum number of surface water specimen to be tested) would improve comparability and interpretability of simulation tests. Harmonisation of test performance can be achieved by providing recommendations in ECHA guidance documents by giving recommendations on typical conditions, but flexibility should be kept to be able to respond to specific cases.

With regard to NER, for cationic substances other extraction procedures than currently discussed organic solvent extractions should be applied (e.g. using high salt solutions, possibly combined with chaotropic agents): due to the ionic interactions significantly contributing to adsorption it is doubtable that organic extraction is effective for these compounds.

7.3 Screening tests

It is recommended that enhanced test modifications should be discussed together with the revision of OECD 301 / OECD 310 tests, but that an own test guideline should be developed for enhanced biodegradation screening tests.

Guidance should be further developed to state more precisely as to where when simulation tests (e.g. for identifying degradation products) are required for CSA. Different data may be required for P assessment compared to risk characterisation.

It was noted that volatile substances are difficult to test in simulation tests. This is considered a topic for further research.

Annex I – Agenda

Tuesday 16th February 2016

12:00 – 12:50	Arrival - possibility to take lunch at UBA canteen	Meeting point: Conference room 0.163
13:00 – 13:15	Address of Welcome	Prof. Dr. Ing. Adolf Eisenträger (Head of the Department)
13:15 – 13:30	Introduction to the topic and the literature study	Klaus Schneider, FoBiG
13:30 – 13:50	Background: Regulatory requirements for assessment of persistence	Markus Schwarz, FoBiG
13:50-14:05	Discussion	
14:05 – 14:25	Screening tests for ready biodegradation: deficits and possible improvements	Stefan Gartiser, Hydrotox GmbH
14:25 – 14:40	Discussion	
14:40 – 14:50	Break	
14:50 – 15:10	Screening tests for enhanced biodegradation: current developments	Stefan Gartiser, Hydrotox GmbH
15:10 – 15:25	Discussion	
15:25 – 16:00	Coffee break	
16:00 – 18:00	Round table discussion: on (enhanced) screening tests for ready biodegradation	
19:30	Joint dinner	*Brauhaus zum alten Dessauer*, Lange Gasse 16

Wednesday 17th February 2016

9:00 – 9:15	Wrap-up of day 1	Klaus Schneider, FoBiG
9:15 – 9:35	Compartment-specific screening tests: recent developments	Thomas Junker, ECT GmbH
9:35 – 9:50	Discussion	
9:50 – 10:10	Simulation tests: current status and outlook	Thomas Junker, ECT GmbH
10:10 – 10:25	Discussion	
10:25 – 10:55	Coffee break	
10:55 – 12:45	Round table discussion: on • compartment-specific screening tests, • simulation tests • research needs and generic strategies for persistence evaluation	
12:45 – 13:00	Wrap-up	Klaus Schneider, FoBiG
13:00	Possibility to take lunch at UBA canteen, departure	

Annex II – Participants of the workshop

The list of participants is attached to this document.

Name	Institution
Ackermann, Juliane	German Environment Agency
Balejíková, Jana	Centre for Chemical Substances and Preparations of Ministry of Economy
Bintein, Sylvain	European Commission, DG Environment
Bonnomet, Vincent	European Chemicals Agency
Claßen, Daniela	German Environment Agency
Doyle, Ian	Environment Agency (UK)
Einola, Juha	Finnish Safety and Chemicals Agency
Gartiser, Stefan	Hydrotox GmbH
Jöhncke, Ulrich	German Environment Agency
Junker, Thomas	ECT Ökotoxikologie GmbH
Knäbel, Anja	Swiss Federal Institute of Technology in Zurich (ETH Zürich)
Mendonça, Elsa	Portuguese Environment Agency
Rauert, Caren	German Environment Agency
Schmidt, Jana	German Environment Agency
Schneider, Klaus	FoBiG GmbH
Schwarz, Markus	FoBiG GmbH
Staude, Claudia	German Environment Agency
Straczek, Anne	French Agency for Food, Environmental and Occupational Health Safety (ANSES)
Tyle, Henrik	Danish Environment Protection Agency
van der Jagt, Katinka	European Commission, DG Environment
Verbruggen, Eric	National Institute for Public Health & Environment (RIVM)

8 References

Baginska, E.; Haiß, A.; Kümmerer, K. (2015)

Biodegradation screening of chemicals in an artificial matrix simulating the water-sediment interface

Chemosphere, 119, 1240-1246

Junker, T.; Coors, A.; Schüürmann, G. (2016)

Development and application of screening tools for biodegradation in water - sediment systems and soil

Science of the Total Environment, 544, 1020-1030

Junker, T.; Paatzsch, C.; Knacker, T. (2010)

A water-sediment screening tool for measuring biodegradation of organic chemicals

Science of the Total Environment, 408, 3803-3810

Rauert, C.; Friesen, A.; Hermann, G.; Jöhncke, U.; Kehrer, A.; Neumann, M.; Prutz, I.; Schönfeld, J.; Wiemann, A.; Willhaus, K.; Wöltjen, J.; Duquesne, S. (2014)

Proposal for a harmonised PBT identification across different regulatory frameworks

Environmental Sciences Europe, 26:9, 1-12