Technical Report - 2014 - 077

TECHNICAL REPORT ON AQUATIC EFFECT-BASED MONITORING TOOLS

Environment

Europe Direct is a service to help you find answers
to your questions about the European Union

New freephone number:
00 800 6 7 8 9 10 11

A great deal of additional information on the European Union is available on the Internet.
It can be accessed through the Europa server (http://ec.europa.eu).
Luxembourg: Office for Official Publications of the European Communities, 2014

ISBN 978-92-79-35787-9
doi: 10.2779/7260

TECHNICAL REPORT ON AQUATIC EFFECT-BASED MONITORING TOOLS

Activity Leaders:

Ann-Sofie Wernersson
(Swedish Agency for Marine and Water Management, Sweden) - Chair
Mario Carere (ISS-Italian Institute of Health, Italy)
Chiara Maggi (ISPRA- Institute for Environmental Protection and Research, Italy)

Drafting Group:

Petr Tusil, Premysl Soldan (T.G. Masaryk Water Research Institute, Czech Republic)

Alice James, Wilfried Sanchez (INERIS, France)

Katja Broeg, Ulrike Kammann (Thünen Institute of Fisheries Ecology), Georg Reifferscheid, Sebastian Buchinger (Federal Institute of Hydrology) (Germany)

Hannie Maas (Institute of Water, Transport and Management), Esther Van Der Grinten (RIVM), (The Netherlands)

Simon O'Toole (EPA-Ireland)

Antonella Ausili, Loredana Manfra (ISPRA- Institute for Environmental Protection and Research), Laura Marziali, Stefano Polesello (IRSA-CNR), Ines Lacchetti, Laura Mancini (ISS-Italian Institute of Health) (Italy)

Karl Lilja, Maria Linderoth, Tove Lundeberg (Swedish Environmental Protection Agency), Bengt Fjällborg, Tobias Porsbring (Swedish Agency for Marine and Water Management), Joakim Larsson, Johan Bengtsson-Palme, Lars Förlin (University of Gothenburg) (Sweden)

Robert Kase, Cornelia Kienle, Petra Kunz, Etienne Vermeirssen, Inge Werner (Eawag/EPFL, Switzerland)

Craig D. Robinson (Marine Scotland Science, Scotland)
Brett Lyons, Ioanna Katsiadaki (Cefas), Caroline Whalley (Cefas) (UK)

Klaas den Haan (CONCAWE)

Marlies Messiaen (Eurometaux)

Helen Clayton (DG Environment, European Commission)

Teresa Lettieri, Raquel Negrão Carvalho, Bernd Manfred Gawlik (European Commission, DG Joint Research Centre)

Valeria Dulio (INERIS-Norman Network, France)

Henner Hollert, Carolina Di Paolo (RWTH Aachen University-Norman Network, Germany)

Werner Brack (UFZ-Norman Network, Germany)

CONTENTS

SUMMARY

BACKGROUND TO THE REPORT

The use of effect-based tools has been mentioned in the context of the Water Framework Directive (WFD) in the Common Implementation Strategy (CIS) guidance no.19 (on water chemical monitoring), in the CIS guidance no. 25 (on sediment and biota monitoring), and (in relation to sediment assessment) in the CIS guidance no. 27 (on environmental quality standards).

In the mandate for 2010-2012 of the sub-group CMEP (Chemical Monitoring and Emerging Pollutants) of the Working Group on Chemical Aspects under the CIS for the WFD, a specific task or activity (3.2 C) was foreseen for the elaboration of a technical report on effect-based tools. The activity was chaired by Sweden and co-chaired by Italy and progressively involved several Member States and stakeholders in an EU-wide drafting group supported by several additional European scientific experts in the field of effect-based tools.

Two drafting group meetings were organised at the Oekotoxzentrum of the EAWAG Institute in Dübendorf in Switzerland on 8 November 2011 and 8 May 2012.The aim of the report, as written in the mandate for the CMEP, was *to identify potential effect-based tools (e.g. biomarker, bioassays) that could be used in the context of the different monitoring programmes (surveillance, operational and investigative) linking the chemical and ecological status assessment*.

The report was approved by the CMEP sub-group in Gent (October 2012), by the WG on Chemical Aspects in Bruxelles (April 2013), by the SCG in Bruxelles (October 2013) and endorsed by the Water Director Meeting in Vilnius (December 2013)

WHY MONITOR EFFECTS?

Chemical analysis generally requires a priori knowledge about the type of substances to be monitored whereas, for technical and economic reasons, it is not possible to analyse, detect and quantify all substances that are present in the aquatic environment. Chemical monitoring is therefore usually focused on already regulated substances that are known to pose a threat to or via the aquatic environment. There is a need to understand which are the real effects caused by the sum of the chemical substances in the aquatic environment (including emerging pollutants, metabolites and transformation products) and to link the observed effects with cost-effective management objectives. Furthermore, the substances present in the aquatic environment can form mixtures whose effects may not be predictable on the basis of chemical analyses alone.The key advantage of monitoring effects is that the overall response from co-exposure to multiple, bioavailable chemicals can be taken into account, including on different levels of biological organisation, such as community, population, individual and/or suborganism level.

In the context of the WFD, the use of effect-based tools can be foreseen for the elaboration and implementation of monitoring programmes and could be used to support the assessment of water quality and provide a link between chemical and ecological assessments. The main aim of this technical report is to present the state of the art of aquatic effect-based monitoring tools for toxic substances from a broader WFD perspective, and to describe in which way these tools can help Member States to rationalise monitoring programmes (including to reduce monitoring costs). The report also contains specific sections on the use of such tools in marine contexts such as the Regional Seas Conventions and the Marine Strategy Framework Directive (MSFD). The MSFD has foreseen the use of effect-based tools; in particular the indicators related to descriptor 8 of the MSFD should also include effects from hazardous substances on ecosystem components.

THE TOOLBOX

For simplicity, the tools are categorised in the report into three main groups, primarily depending on the type of monitoring approach used:

1) Bioassays, *in vitro* and *in vivo*, which measure the toxicity of environmental samples under defined laboratory conditions, on cellular and individual levels respectively.

2) Biomarkers, i.e. biological responses at individual level or below, observed in field exposed organisms

3) Ecological indicators, that measure variations observed at higher biological organisation levels, i.e. population and community.

Several of the tools described in this report are already used also outside the scientific community in Europe, for both marine and limnic applications. Biomarkers are included in the monitoring programmes of Regional Seas Conventions to identify the presence of substances or combinations of substances not previously identified as a concern and to identify regions of decreased environmental quality. Bioassays are used for example to support the risk assessment and management of contaminated sediment and to provide decision support to prohibit the release of toxic substances into the environment (e.g. in the evaluation of dredged sediments that are considered for sea disposal and Whole Effluent Assessments in the permitting process). They are also used in the broad screening of different sources (such as sewage treatment plant effluents). Other applications include for example alarm systems, which directly trigger control measures (closing drinking water intakes) Based on these experiences, it can be concluded that the main objectives of using effect based monitoring tools within the current WFD context would be:

- As screening tools, as part of the pressures and impacts assessment to aid in the prioritisation of water bodies to study further.

- To establish early warning systems, to prioritise further studies in areas that are not concluded to be at risk because they are located far from known local sources.

- To take the effects from mixtures of pollutants or not analysed chemicals into account (e.g. to support investigative monitoring where causes of a decline of specific species are unknown).

- To provide additional support in water and sediment quality assessment, though not as a replacement for conventional chemical and ecological monitoring under the WFD.

The technical report includes a dedicated section on the use of effect-based tools in the different Member States and in the context of the Regional Seas Conventions (e.g. OSPAR). The report also includes descriptions of tools and methodologies that are considered promising in the near future because of the fast development in this area. The identity of the substances causing the main response in a bioassay can be investigated using, for example, a novel approach based on fractionation techniques ("Effects Directed Analysis"- EDA), and community-level effects that are specifically linked to certain substances can be assessed by using certain ecological indicators, such as SPEAR. Furthermore in this report there is a specific section related to recent research development in OMICs technologies that could have a potentially wide future application in the monitoring and assessment of aquatic environments.

The report has an Annex that contains 14 case studies, illustrating how these tools can help to achieve the objectives of the WFD and MSFD; a series of fact-sheets providing technical specifications for some relevant aspects to consider in the design of aquatic monitoring programmes for selected individual effect-based tools (biomarkers and bioassays) that are either already used on a routine basis or are gaining in popularity. The Annex contains also a list of available standardised effect-based tools (*in vivo* and *in vitro* bioassays and biomarkers), established assessment criteria for the marine environment and an overview of available DNA microarrays. Also some technical issues, such as sampling aspects, standardisation issues and proposed approaches to assess oestrogenic effects are described in more detail in the Annex. Finally, a list of definitions, abbreviations and a reference section are included.

OUTLOOK

The topic "effect based tools" ranked high in the CIS science-policy interface report elaborated on the basis of inputs from the WG on Chemical Aspects. The new mandate 2013-2015 of the WG Chemicals of the WFD, approved by the Water Directors, has decided that work on effect-based tools should continue, in particular in relation to the detection and evaluation of effects caused by mixtures of pollutants. This activity will be strongly linked to the work of the WG Ecostat and the implementation needs of the MSFD.

This technical report, elaborated in close collaboration with the scientific community, can already be considered to provide important support to the managers, the assessors and the local operators involved in the analysis and monitoring of surface water.

1 INTRODUCTION

1.1 Why a technical report on effect-based monitoring tools is needed

The Water Framework Directive (WFD), 2000/60/EC, requires an integrated approach to the monitoring and assessment of the quality of surface water bodies. The assessment of ecological status takes account of effects at population and community level, based on the use of specific indices and ecological quality ratios. The chemical status assessment is based on compliance with legally binding Environmental Quality Standards (EQSs) for selected chemical pollutants (priority substances) of EU-wide concern. EQS for priority substances are set in the Directive 2008/105/EC, recently amended by 2013/39/EU, while EQS for other pollutants are set by Member States for substances of national concern assessed under ecological status; the EQSs are designed to protect the environment and human health (see also chapter 2.3 on regulatory background). Thus, chemical analysis is a fundamental step in the assessment of surface water status under the WFD. In the context of the MSFD (Marine Strategy Framework Directive; 2008/56/EC), monitoring requirements relating to contaminants cover their biological effects as well as direct chemical analysis.

Chemical analysis generally requires *a priori* knowledge about the type of substances to be monitored whereas, for technical and economic reasons, it is not possible to analyse, detect and quantify all substances that are present in the aquatic environment, in all waters. There is therefore a tendency to focus chemical monitoring on already regulated substances that are known to pose a threat to or via the aquatic environment. Even for the 7336 unique substances registered under REACH in the year 2010 it would be challenging if not impossible to plan a chemical monitoring programme on all that could pose a risk to the aquatic environment. Furthermore, to estimate the risk of effects related to the large number of substances that are present and detected in the environment (including emerging pollutants, metabolites and transformation products), it would be necessary to develop a very large number of assessment criteria (quality standards). Such assessment criteria for chemicals are generally developed substance by substance, based on laboratory studies, and usually do not consider the consequences of the co-exposure to multiple chemicals that occurs in the environment, possibly giving rise to cumulative effects (see e.g. Silva et al 2002). The communication of the European Commission on combination effects of chemicals (COM 2012-252) also suggests that the Commission should initiate a work programme to ensure that risks associated with chemical mixtures are properly understood and assessed. The report states that EU laws set strict limits for the amounts of particular chemicals allowed in food, water, air and manufactured products, but that the potentially toxic effects of these chemicals in combination are rarely examined. Another aspect to be considered is that chemical and ecological assessments give often different results. For management purposes, there is a need to understand which are the real effects caused by chemical substances in the aquatic environments, and to link observed effects with cost-effective management objectives.

Aquatic effect-based monitoring tools, developed to respond to toxic substances, offer possibilities to overcome some of the limitations. The key advantage with monitoring also biological effects is that the overall response from co-exposure to multiple, bioavailable chemicals can be taken into account, at different levels of biological organisation, such as community, population, individual and/or suborganism levels. In this way, a more holistic approach is possible.

Whereas the WFD chemical status compliance assessment should be based on concentrations of listed substances, the aim of this technical report is to present the state of the art of aquatic effect-based monitoring tools for toxic substances from a broader WFD perspective, and to describe in which way these tools can help Member States to rationalise monitoring programmes (including to reduce monitoring costs). The tools described in this report are therefore primarily related to the protection of the aquatic environment from direct exposures and focused on fish, invertebrates and other aquatic biota[1]. Different monitoring approaches that can be used to detect and assess effects from hazardous substances on aquatic ecosystems or components thereof are described.

1.2 Categories of effect-based monitoring tools

In this report, effect-based monitoring tools are categorised into three main groups, primarily depending on the type of monitoring approach used (rather than the type of effect studied). A review of these major categories of tools, within both risk assessment of chemicals and aquatic monitoring, has recently been published (Connon et al 2012). Please note that the definitions used below may deviate from definitions used elsewhere. A list of terms and defintions used in this report is therefore included in the Annex (section 12).

1) Bioassays, *in vitro* and *in vivo*, which measure the toxicity of environmental samples under defined laboratory conditions (chapter 4)

2) Biomarkers, i.e. biological responses at individual level or lower organisational levels, observed in field-exposed organisms (chapter 4)

3) Ecological indicators, that measure variations observed at higher biological organisational levels, i.e. population and community (chapter 5)

In vitro bioassays are based mainly on cell lines (lower biological organisational level), responding to those compounds in a sample that have the same mode of action, such as binding to a specific cellular receptor or change in a specific DNA component. They have much in common with chemical analytical screening tools, but a “biological detector” is used and therefore these bioassays are often refered to also as “bioanalytical tools”. More or less any type of sample can be analysed, and the results are frequently expressed on a chemical equivalent basis. However, they measure the cumulative effect from all substances in the sample having the same mode of action and not only that particular

[1] Some of the tools can also be used for an early evaluation of possible risks for human health (e.g. mutagenicity test), in relation to consumption of fishery products or drinking waters.

substance. In addition, they can integrate and quantify antagonistic modes of action that could potentially be present in the sample. Because of the comparatively low costs and ability to detect a large number of substances having the same mode of action, several *in vitro* bioassays are suitable for screening purposes. They are also usually suitable for high throughput applications.

In vivo bioassays[2] are performed using live organisms. They have the capacity to provide an integrated response at organism level to contaminants[3] in a sample. In general, highly relevant endpoints (such as survival, growth and mobility) are analysed.

Biomarkers can be used to study effects such as biochemical, physiological, histological, or morphological alterations in field exposed[4] individuals. They are sometimes divided into *specific* and *general* biomarkers. The latter respond to several types of substances and possibly also other stressors than hazardous substances. Specific biomarkers are generally related to a limited number of substances. Another categorisation is into *exposure* and *effect* biomarkers, referring to the ecological relevance of the endpoint analysed. There is no sharp line between these categories, but exposure biomarkers are considered very useful as early warning signals whilst effect biomarkers can be used for ecosystem risk assessment; specific biomarkers can more easily be related to a particular pressure whilst general biomarkers have the capacity to integrate the response related to several stress factors and thus also toxicologically induced responses from contaminant mixtures.

Ecological indicators are related to the impact on community structure and/or function, and generally provide a highly integrated and relevant response. It can be difficult to identify underlying causes, because the response can be due to combined effects of several types of stressors. However, in chapter 5, a few promising indicators that also can be linked to chemical stressors are presented. In a WFD context, these indicators therefore show potential to be useful in the future as biological quality elements within ecological status classification.

The identity of the substances causing the main response in a bioassay can be investigated using a novel approach based on fractionation techniques ("Effects Directed Analysis", EDA, further described in chapter 6).

Furthermore in this report there is a specific section related to recent research development in the field of OMICs technologies that could have a potentially wide future application in the monitoring and assessment of aquatic environments.

1.2.1 Considerations regarding duration and effects of exposure

The ability of any biological method to indicate the state of the environment is dependent on the degree and duration of exposure to the pollutants, as well as the sensitivity and response rate of biological processes, see fig 1. The effects demonstrated by organisms experiencing environmental stress from chemical exposure range from biochemical changes at the sub-cellular level to death or migration from the affected area. Many of

[2] They are frequently performed in more or less the same way as within chemicals testing, but then normally called "toxicity tests".
[3] If performed under standardised laboratory conditions, other factors than chemical can normally be ruled out.
[4] Normally caught from the wild, but sometimes cultured organisms are exposed in the field (cage experiments)

these effects have been incorporated or developed into possible monitoring methods, but most are only applicable to certain types of environmental stress, to particular time scales of stress effects or to specific habitats or localities.

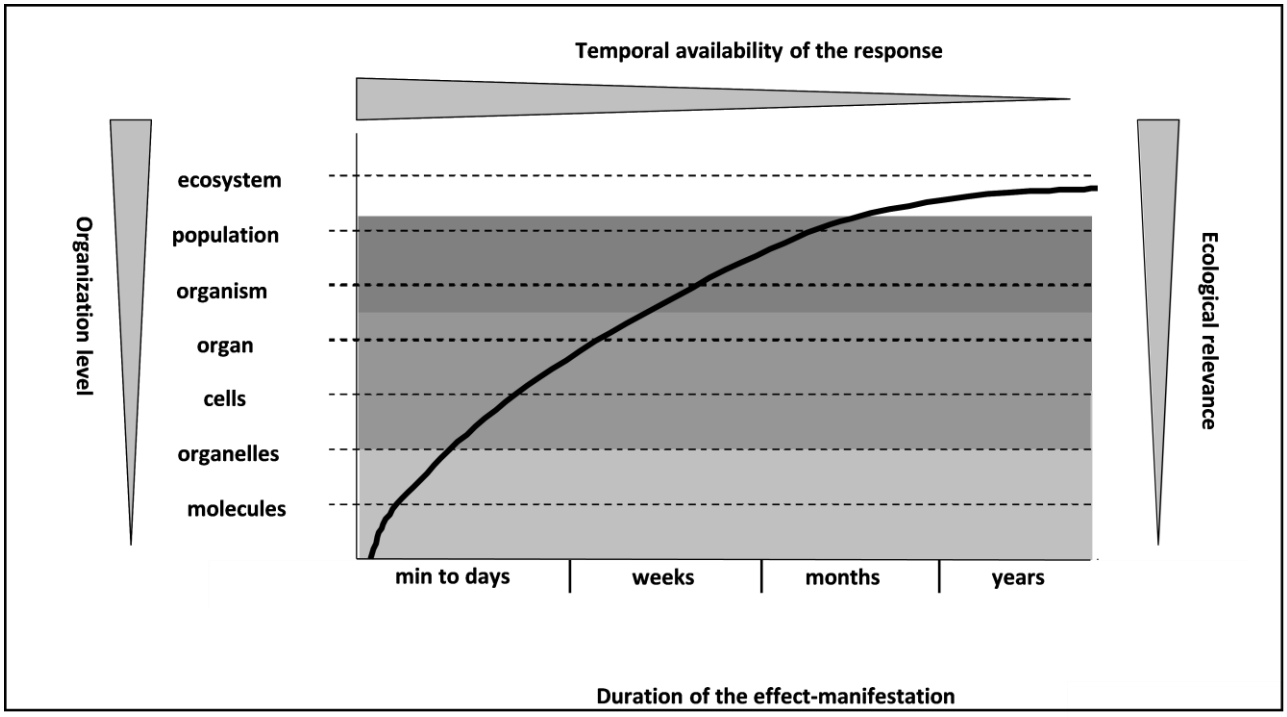


Fig 1. Response time and ecological relevance at different levels of biological organisation

Effect-based monitoring tools at lower organisational levels can, like chemical analysis, be used to predict relevant effects, but both approaches should preferably be performed and assessed in an integrated manner. For ethical, scientific, practical and economic reasons, it is not possible to monitor all types of effects directly, and the prediction of long-term risks from substances that have biomagnification potential usually requires chemical analysis.

1.3 Structure of the report

Chapter 2 includes a description of the current use of effect-based tools in the regulatory context such as under the Regional Seas Conventions and in relation to the implementation of the MSFD, as well as by individual Member States. Chapter 3 proposes, in general terms, how effect-based tools described in this report could be used within WFD and MSFD monitoring contexts.

Technical and scientific aspects of available tools and monitoring approaches are described in chapter 4 (biomarkers and bioassays) and 5 (ecological indicators).

Additional, novel tools and approaches are described in chapter 6 (Effects Directed Analysis, EDA) and chapter 7 (OMICS).

The report has an Annex that contains case studies, illustrating how these tools can help to achieve the objectives of the WFD and MSFD; a series of fact-sheets providing technical specifications for some relevant aspects to consider in the design of aquatic monitoring programmes for a selected number of individual effect-based tools (biomarkers and bioassays) that are either already used on a routine basis or are gaining in popularity; a list of available standardised effect-based tools, established assessment criteria for the marine environment and an overview of available DNA microarrays. Also some technical issues, such as sampling aspects, standardisation issues and proposed approaches to assess oestrogenic effects are described in more detail in the Annex. Finally, a list of definitions and abbreviations and a reference section are included.

2 CURRENT USE AND REGULATORY RELEVANCE OF EFFECT-BASED MONITORING TOOLS

The usefulness of aquatic effect-based monitoring tools is supported by extensive scientific literature, national and European research projects, such as MODELKEY[5], Balcofish[6] and EDA-Emerge[7] and Regional Seas Conventions. Some effect-based tools are applied by national and regional authorities in different contexts. In a WFD context, the assessment of ecological status takes account of effects at population and community level, whereas indicators related to descriptor 8 of the MSFD should also include effects of hazardous substances on ecosystem components (including individuals).

2.1 MSFD and Regional Seas Conventions

In the context of the MSFD, monitoring of the biological effects of contaminants is required. Commission Decision 2010/477/EU specifies criteria and methodological standards for evaluating whether Good Environmental Status (GES) is achieved. According to the Decision, effect-based indicators are to be included to assess GES, related mainly to descriptor 8, "Concentrations of contaminants are at levels not giving rise to pollution effects". Within the MSFD context, harmonisation between marine regions, and the Regional Seas Conventions as well as with the assessments made under the WFD is emphasised.

Integration of the results of chemical monitoring programmes, and combination of data from chemical and effect-based monitoring, is already an active area of science within the Regional Seas Conventions. Integrated monitoring programmes, data management, assessment, interpretation and presentation schemes are being developed and/or already applied. Current experience indicates that integration is greatly facilitated by coherent and consistent sets of assessment thresholds (such as Environmental Assessment Criteria (EACs) and Background Assessment Concentrations (BACs); see also Annex section 8).
Furthermore, The International Council for the Exploration of the Sea (ICES) has produced detailed reports (ICES, 2011; Davies & Vethaak 2012) on integrated monitoring of contamintants and their effects which have been accepted by OSPAR for use within the Joint Assessment and Monitoring Programme (JAMP), on a three-year trial basis. The scheme describes the use of assessment criteria indicative of background conditions and of "harm" to assess environmental status from data on a range of biomarkers, higher-level effects (e.g. disease, biossays) and chemical concentrations in biota (fish, mussels, gastropods), water, and sediments. Information on benthic faunal invertebrate communities could also be incorporated into the approach. The different components of this integrated approach are listed in table 4.2. (chapter 4) of this report.

The CEMP (Coordinated Environmental Monitoring Programme) is already mandatory for all contracting parties of OSPAR and includes one effect-based tool[8] to evaluate one of the

[5] http://www.modelkey.org/
[6] http://www.balcofish.science.gu.se/english
[7] http://www.ufz.de/eda-emerge/
[8] Imposex biomarker

Ecological Quality Objectives[9]). There are also several "preCEMP" effect-based tools that are not mandatory but recommended and included in the agreements made by several of the individual contracting parties. Such components include both biomarkers and bioassays[10]. In June 2013, the OSPAR Commission agreed to adopt a set of "common indicators", to become components of the OSPAR monitoring and assessment work, in particular contributing to the OSPAR Intermediate Assessment in 2017. This assessment is intended to contribute, for those Contracting Parties that are also bound by the MSFD, to the updating of their Art. 8 assessment under that Directive (due in 2018). The common indicators are also intended to facilitate achieving (sub)regional compatibility of monitoring programmes as required by MSFD Art. 11(1). Also a set of candidate indicators have been identified. For such indicators, further development is required before a decision can be taken to either adopt each of them as a common indicator or to remove them from the list. Common and candidate indicators are listed in table 4.2. (chapter 4).

In a similar manner, core variables are mandatory monitoring components within HELCOM. The project HELCOM CORESET was initiated to identifiy suitable core variables to be used both to evaluate progress in reaching the Ecological Objectives[11] related to the Baltic Sea Action Plan and as indicators related to the MSFD for those contracting parties that are also European Member States. In an interim report, six core indicators[12] and three candidate indicators[13] related to effects of hazardous substances were proposed (HELCOM 2012a, b). According to a decision made at HELCOM MONAS meeting (18/2013), the aim is that the set of core indicators will be measured by all contracting parties and that contracting parties that are also EU member states will use the core indicators for the MSFD implementation. Pre-core indicators will be under further development under the HELCOM CORESET II project (during 2013-2015) with the aim to include them in the core set by HELCOM HOD in 2015. In addition there are also candidate indicators (to be developed into a core indicator proposal) and supplementary indicators (applied on a sub-regional basis). The finally decided effect-based core, precore and candidate indicators of HELCOM are listed in table 4.2. (chapter 4). Under the MEDPOL Convention, a lysosomal stability test has been proposed as a biomarker in pilot studies.

2.2 Current use of aquatic effect-based monitoring tools in Europe

At national level in Europe, the traditional chemical approach has been combined with monitoring tools that analyse the effects of hazardous substances also in contexts other

[9] OSPAR EcoQ related to TBT effects: "The average level of imposex in a sample of not less than 10 female dogwhelks (Nucella lapillus) should be consistent with exposure to TBT concentrations below the environmental assessment criterion for TBT. Where Nucella lapillus does not occur naturally or where it has become extinct, other species may be used."

[10] CYP 1A activity (EROD), bulky aromatic-DNA adducts, PAH metabolites in bile, liver histopathology, macroscopic liver neoplasms (= liver nodules), ALA-D, Metallothionein (MT), sediment and water bioassays, lysosomal stability, externally visible fish diseases and reproductive success in eelpout.

[11] One of the EcoOs is "Toxic substances shall not cause sublethal, intergenerational or transgenic effects to the health of marine organisms (e.g. reproductive disturbances)."

[12] Imposex biomarkers, lysosomal membrane stability, fish disease index, micronucleus induction, malformed embryos of amphipods, malformed embryos of eelpout

[13] Vitellogenin induction, Acetylcholine Esterase, EROD/CYP1A

than the Regional Seas Conventions. Effect-based tools are sometimes required by national and regional authorities, for example in investigations of dredged sediment and at contaminated sites as well as within Whole Effluent Assessments (WEA), see e.g. OSPAR guidance (2007) and COHIBA (2010). Current use in some European countries is briefly described below, primarily focusing on use other than in research. It is not a complete overview and additional applications may exist.

In the Czech Republik, ecotoxicological tools are not required for monitoring, but they are frequently used to detect hazards in the aquatic environment, on both a large geographical scale and for targeted studies. They involve a wide range of different approaches, including *in vivo* and *in vitro* tools. See also case study « The risk of chronic impact of pollution on the Bilina river ».

Denmark has applied effect-based tools for several years for both monitoring and assessment purposes. Imposex biomarkers in four species of whelks (the neogastropods: *Neptnunea antiqua, Buccinum undatum, Nucella lapillus* and *Hinia reticulate*) and intersex in periwinkle (*Littorina littorea*) have been a part of the national monitoring programme since 1998; fry abnormalities and enzyme activity (EROD) in eelpout (*Zoarces viviparus*), and lysosomal membrane stability in haemocytes of mussels (*Mytilus edulis*) since 2004. A new addition in 2011 to the national monitoring programme is PAH metabolites in bile from eelpout. Another part of the national monitoring programme is monitoring campaigns regarding effects of endocrine disrupting substances like intersex and sex ratio in eelpout and use of *in vitro* assays for assessing oestrogenic activity. Various research projects have also included studies with other types of biomarkers[14] in mussels and fish in Danish waters. These studies have made use of the current activities in the national monitoring programme such as coordinated sampling.

For the assessment of dredged material in Germany, the use of effect-based tools is foreseen. For the Federal waterways, the Federal Ministry of Transport (Bundesministerium für Verkehr BMV) and its subordinate authorities are responsible. Conceptual guidance and project monitoring with regard to environmental aspects are covered by the Federal Institute of Hydrology (Bundesanstalt für Gewässerkunde BfG). All other inland waterways are under the responsibility of the Länder (Federal states) which have their own guidelines and recommendations (den Besten et al. 2003). For the Federal waterways two directives apply: the Directive for the Handling of Dredged Material on Federal Inland Waterways for the freshwater area (HABAB-WSV 2000, Bundesanstalt für Gewässerkunde), which specifies an algae test, luminescent bacteria test, and daphnia test , and the Joint Transitional Arrangements for the Handling of Dredged Material in German Federal Coastal Waterways (GÜBAK-WSV 2009[15]) which specifies a marine algae test, luminescent bacteria test, and amphipod test. Additonally, in case-by-case studies, a broad set of *in vitro* bioassays is applied for evaluating German surface waters, suspended particulate matter and sediments. Reviews are given by Hollert et al. (2009) and Hallare et al. (2011).

In Italy effect-based tools are used for sediment quality assessment in marine and transitional coastal waters when sediment EQS are exceeded in compliance with the recommendations of CIS guidance no. 27. The tools used are *in vivo* – acute and chronic

[14] The studies include several biomarkers (such as immune response in mussels, clams and fish micronuclei, DNA damage, AChE activity in mussels and fish, embryo aberrations in amphipods, vitellogenine induction in fish, ALP in mussels, etc).

[15] http://www.bafg.de/Baggergut/DE/04_Richtlinien/guebag_en.pdf?__blob=publicationFile

bioassays using up to 3 trophic levels (see technical guidance *Manuale per la movimentazione dei sedimenti marini"* ICRAM-APAT 2007). Effect based tools are used also to support the WFD investigative monitoring programmes (*in vivo*-bioassays, biomarkers) (Ministerial Decree 260/2010) and required as part of the authorisation for dredged sediment disposal at sea and for the evaluation of discharges of urban and industrial effluent (*Daphnia magnia* test). Effect based tools are also used for the waste characterization and classification.

In Ireland, effects-based tools are not widely used for regulatory or environmental monitoring with some limited exceptions, *e.g.* licensing of certain pharmaceutical plants. An EPA-funded research project carried out a number of years ago (Tarrant et al, 2005) on the effects of endocrine disruptors in the Irish environment found levels in the low ng/L range, as determined by the YES assay. Results were supported by separate measurements of vitellogenin levels in male wild Brown trout. See also the case study "Endocrine disruptors in the Irish Aquatic Environment".

The Netherlands has applied effect-based tools for many years for various monitoring and assessment purposes. Application of these effect-based tools, however, has never been prescribed in the national law. Implementation of the WFD and the fact that bioassays are not obliged in monitoring programmes has led to a re-evaluation of the use of these tools. Due to policy and economic reasons only a few applications are therefore still required on occasion. Until 2010, *in vivo* bioassays on concentrated water extracts were included in freshwater monitoring programmes on a regular basis (see case study "Monitoring concentrated surface water with *in vivo* bioassays"). Still in use is a combined application of on-line continuous biotests (*Daphnia magna* and algae) and advanced chemical analysis in the main rivers (Meuse and Rhine) to detect unknown hazardous substances arising from accidents. Also drinking water suppliers make use of these systems, and a positive alarm triggers immediate action (closing intake or adjusting treatment) as well as identification procedures (see e.g. De Hoogh et al 2006). Imposex biomarkers in gastropods are included in the marine national monitoring programme and considered for the marine monitoring strategy under the MSFD. Effect-based tools are occasionally required in the assessment of contaminated sediments, but applied only if a chemical compliance check and ecological assessment of the water body give reason for additional assessment. Additional, effect-based tools are occasionally included in screening projects or in projects to characterise sources. These tools include DR- and ER-CALUX to monitor dioxins and oestrogens and tests for antiobiotics or compounds that cause DNA damage (Comet assays).

In Sweden, the marine national monitoring programmes include several biomarkers, such as imposex in gastropods, marine bird egg shell thinning, Monoporeia embryo deformations and an integrated fish monitoring programme (population level effects, tissue concentrations of persistent substances, and health-related variables), to fulfil OSPAR and HELCOM requirements, but also to support the evaluation of progress towards reaching the national environmental quality objectives[16]. The fish health programme includes a large biomarker battery and also other variables than those required or recommended within the marine regional conventions (see case study "Swedish national monitoring programme of fish health"). The purpose is primarily to observe trends and to act as early warning of large-scale negative effects from hazardous substances. A few so-called coordinated

[16] Three of the sixteen objectives are "A non toxic environment", "Flourishing lakes and streams" and "A balanced marine environment"

impact monitoring programmes[17] can also include biomarkers, either on a regular basis or in a campaign-wise manner. The annual national screening campaigns occasionally so far also included bioassays to characterise certain sources and the interest in these screening tools is increasing. Oestrogenic and androgenic effects of treated effluents from selected sewage treatment plants are monitored within the annual screening programme. In addition, effect-based tools are on occasion required as part of WEA, primarily during permitting processes for large industrial plants. Such procedures have been described in a recently revised handbook (Naturvårdsverket 2011), and focus is by tradition on *in vivo* bioassays such as *Daphnia magna* and algae, but on occasion also *in vitro* tools are used. Effect-based tools are also sometimes used in investigations of contaminated sediments, to provide additional decision support related to remediation needs. Such investigations have included, e.g. fish biomarkers and menthum deformations in chironomids.

In the UK, CEFAS (Centre for Environment, Fisheries and Aquaculture Science) is developing a weight-of-evidence approach for evaluating dredged sediments for sea disposal, based on physical, chemical and bioassay data in parallel. Until recently, sediment and whole water bioassays were used to assess UK coastal and transitional waters. A number of fish biomarkers have been routinely used for many years by Cefas and Marine Scotland in relation to OSPAR marine monitoring, particularly in relation to PAHs. Regular imposex surveys in gastropods have demonstrated the effectiveness of the ban on organotin antifoulants. Since 2009, Cefas and Marine Scotland have been trialling the use of integrated monitoring to assess environmental status in line with guidance produced by the Study Group on Integrated Monitoring of Contaminants and Biological Effects (ICES, 2011). Effect based tools[18] are also used within WEA (Environment Agency 2006).

Norway requires bioassays[19] in the assessment of dredged sediment (KLIF 2011b). As part of the OSPAR CEMP programme, fish biomarkers (EROD, CYP1A, blood ALA-D and PAH metabolites in gall bladder) are monitored in cod and imposex biomarkers in dogwhelk (KLIF 2011a).

In Switzerland, effect based tools are not legally required in water or sediment monitoring programmes. However, in the Modular Stepwise Procedure (MSP), a toolbox developed for local authorities to assess water quality in Switzerland, a macroinvertebrate index is included. Steps are underway to extend effect-based tools in the MSP, by including for example bioassays for oestrogens and herbicides.

In conclusion, this overview illustrates that several of the tools described in this report are already used also outside the scientific community in Europe, for both marine and limnic applications. Effect-based tools are used to support the risk assessment and management of contaminated sediment. Biomarkers in particular are included in the monitoring

[17] Such programmes ("Samordnad Recipientkontroll, SRK") are established by operators, including industry, having a potential impact on a particular aquatic environment, in order to fullfil Swedish law, requiring that the operators should have sufficient knowledge about their impact on receiving waters. The monitoring stations therefore include also more impacted areas and limnic environments.

[18] Daphnia magna immobilisation, Freshwater algae inhibition of growth, Oyster Embryo-Larval development, Tisbe battagliai lethality, Marine algae inhibition of growth

[19] *Skeletonema, Tisbe and Crassostrea* bioassays on pore water and DR CALUX on sediment extracts

programmes of Regional Seas Conventions, with the purpose of identifying the presence of substances or combinations of substances not previously identified as a concern, and to identify regions of decreased environmental quality. Bioassays are used to provide decision support to prohibit the release of toxic substances into the environment (e.g. WEA in the permitting process, evaluation of dredged sediments that are considered for sea disposal). They are also used within broad screening of different sources (such as sewage treatment plant effluents). Other applications include the Dutch alarm system that directly triggers control measures (closing drinking water intake). Effect based tools support also the ecotoxicological characterisation and classification of hazardous wastes in the context of the Waste Framework Directive (2008/98/EC). The experience gained from these applications is valuable when considering the possible use of effect-based tools more widely in the context of WFD and MSFD implementation.

2.3 The WFD – brief regulatory background

For Member States an overall aim of the WFD (2000/60/EC) is to achieve “good ecological status” and a “good chemical status” in all bodies of surface water by 2015. Monitoring programmes are required to establish a coherent and comprehensive overview of ecological and chemical surface water status within each river basin district.

2.3.1 Ecological status under the WFD

Good ecological status is defined in Annex V of the WFD in terms of the values of the biological quality elements (phytoplankton, macroalgae, angiosperms, benthic invertebrate fauna, and fish), the hydrological and morphological conditions, and the physico-chemical elements. In combination, the biological quality elements, hydro-morphological and physico-chemical elements determine the classification of each water body. Good ecological status (or potential) requires that the concentrations of the specific pollutants (also called River Basin Specific Pollutants) do not exceed the EQSs set at Member State level in accordance with the procedure laid down in Annex V, section 2.1.6 of the WFD.

Specific pollutants are not “listed” in the same way as the priority substances although there is an indicative, not exhaustive, list in Annex VIII of the WFD, which includes a wide range of substances and groups of substances that can often be detected in surface water bodies. Member States are required to assess whether they or other potential pollutants are discharged in significant quantities into water bodies. In order to assess the ecological status as regards the concentrations of specific pollutants in the water bodies, EQS have to be established at national level.

2.3.2 Chemical status under the WFD and Environmental Quality Standards for priority substances

The WFD requires that good chemical status be achieved by all Member States. The assessment of chemical status is based on monitoring of the list of priority substances in Annex X of the WFD, for which Directive 2008/105/EC (Environmental Quality Standards

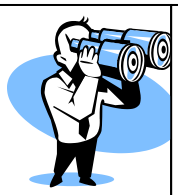
Directive), amended by 2013/39/EU[20], sets standards (EQSs). The Directive sets EQS for particular compartments (water and/or biota) but, under certain conditions, allows Member States to establish EQSs for alternative compartments at national level and to apply those EQSs instead, providing that doing so ensures at least equal protection. The EQSs are concentrations that should not be exceeded, in order to protect human health and the environment. The trends of accumulating priority substances should also be monitored in sediment and/or biota and should not increase (no deterioration objective) in the water bodies.

Recital no. 18 of Directive 2013/39/EU makes explicit reference to the future application of other tools for monitoring: "Novel monitoring methods such as passive sampling and other tools show promise for future application, and their development should therefore be pursued".

The methodology for the derivation of the EQS (at European and national levels) is extensively described in the CIS guidance no. 27.

2.3.3 Analytical requirements

The chemical analysis of priority substances and specific pollutants is subject to stringent quality criteria set out in the Directive on technical specifications for chemical analysis and monitoring of water status (2009/90/EC). All methods of analysis applied by Member States for the purposes of chemical monitoring of water status have to meet certain minimum performance criteria, including rules on the uncertainty of measurements and on the limit of quantification of the methods. Article 4 of the Directive describes the minimum performance criteria for all methods of analysis. However, even for some priority substances, current or proposed EQS values are lower than the current Limits Of Quantification (LOQ), as described in Common Implementation Strategy (CIS) guidance no. 19 and a recent JRC report (Loos, 2012).



Look in:
Directive 2009/90/EC
Loos, 2012

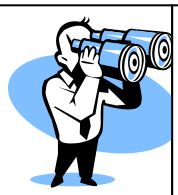
2.3.4 WFD Monitoring Programmes

The WFD requires 3 monitoring programmes

[20] The main features of the revised directive are: 12 additional priority substances, 6 of them designated as priority hazardous substances; stricter EQS for four existing priority substances and slightly revised EQS for three others; the designation of two existing priority substances as priority hazardous substances; the introduction of biota standards for several substances; provisions to improve the efficiency of monitoring and the clarity of reporting with regard to certain substances behaving as ubiquitous persistent, bioaccumulative and toxic (PBT) substances; provision for a watch-list mechanism designed to allow targeted EU-wide monitoring of substances of possible concern to support the prioritisation process in future reviews of the priority substances list; provisions for a strategy on pharmaceuticals.

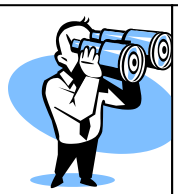
- Surveillance monitoring: to supplement and validate the impacts analysis, to support the efficient and effective design of future monitoring programmes, to assess long-term changes in natural conditions and changes resulting from anthropogenic activity. The monitoring is performed at least once every management cycle (usually every 6 years).
- Operational monitoring: to establish the status of those water bodies identified as being at risk of failing to meet the WFDs environmental objectives, and to assess any changes in the status resulting from the Programme of Measures.
- Investigative monitoring: to determine reasons for exceedances or predicted failure to achieve environmental objectives if the reasons are not already known; and to determine the magnitude and impacts of accidental pollution.

To determine the WFD operational monitoring needs, the analysis of pressures and assessment of impacts is an important first step. In CIS guidance document no. 3 (Analysis of Pressures and Impacts), a pressures and impacts assessment is defined as a four-step process. In the first step the driving forces, such as land-use patterns, are described. In the second step, pressures (point and diffuse souces of hazardous substances) with a potential impact on water bodies are identified. The impacts resulting from the pressures are assessed in a third step. Finally, in the fourth step, the likelihood of failing to meet the objectives is assessed.



Look in:
Directive 2000/60/EC-Annex V- Point 1.3

The hydrophobic substances for their tendency to accumulate must be detected most likely in sediment and biota (see CIS guidance n.25). According to article 3 in the 2008/105/EC directive, Member States shall arrange for the long-term trend analysis of those priority substances that tend to accumulate in sediment and/or biota, giving particular consideration to 20 such substances or group of substances. Biota standards are also expressed for 11 of these substances in the new Directive 2013/39/EC.

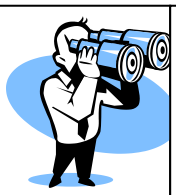


Look at

2008/105/EC revised by **2013/39/EU**
Article 3.6.(trend monitoring)
Annex I column 8 (biota standards)
CIS Guidance n.25-chapter 3, paragraph 5.

2.4 The MSFD – brief regulatory background

The MSFD requires that EU Member States take the necessary measures to achieve or maintain Good Environmental Status (GES) by 2020. Member States need to define GES (Good Environmental Status), set environmental targets as well as establish/revise monitoring programmes, assess the environmental status and establish Programmes of measures in a six-year cycle. MSFD Annex I includes a set of 11 descriptors on the basis of which GES should be determined, and Commission Decision 2010/477/EU[21] includes 29 agreed criteria and 56 indicators on which GES could be defined. Descriptor 8 (D8) and descriptor 9 (D9) refer to concentrations of contaminants. Contaminant monitoring under D8 is highly linked to WFD chemical monitoring, but in addition, contaminant effects are to be monitored and taken into account. D9 refers to chemical contaminants in seafood and is therefore also closely related to D8, but aims to protect human consumers.



COM decision 2010:477

Criteria for D8 and D9

[21] Commission Decision 2010/477/EU of 1 September 2010 on criteria and methodological standards on good environmental status of marine waters.

3 EFFECT-BASED TOOLS IN THE CONTEXT OF THE WFD AND MSFD

3.1 General Aspects - WFD

In the context of the WFD, effect-based monitoring tools are mentioned in some CIS (Common Implementation Strategy) guidance documents: numbers 19 , 25 and (briefly) 27 (in relation to sediment).

In CIS guidance no. 19 (chapter on complementary methods) it is, for example, stated that "it is desirable to introduce other techniques for improving the quality of the assessment and to benefit from resource saving developments, as they become available". In this context, effect-based tools are mentioned (referred to as "Biological assessment techniques"), but it is also stated that these tools "…are designed to respond to a wide range of (chemical) stressors, and are therefore, not exclusively linked to individual quality elements such as the different priority substances. Although very useful for many monitoring purposes, they cannot be used to check compliance of individual quality elements against an EQS." Nevertheless, the following objectives were identified:

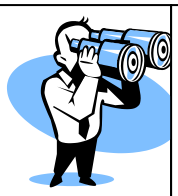
- Early detection of biological imbalance
- Linking ecological and chemical information
- Linking concentration with exposure and effects
- Early warning of changes in water quality at crucial sites
- Detecting and assessing significant pollutants to update risk assessments
- Detecting adverse biological effects to indicate where operational or investigative monitoring is required

In an investigative monitoring context, effect-based tools are suggested as part of the investigations in the event of MAC-EQS exceedence, but also as an early warning to help identify compounds in future risk assessments. By combining passive sampling and effect-based tools, an integration of exposure and effects monitoring is considered to "facilitate more cost effective monitoring programmes as well as forming the basis of a risk-based pollution control strategy." EDA/TIE is mentioned as a possible way to select other compounds (River Basin Specific Pollutants) based on ecological relevance, and effect-based tools "may be able to provide additional weight-of-evidence, mostly in cases where additional information on chemical quality or links between chemical and biological data is required".

In CIS guidance no. 25 (chapter on Complementary Methods), effect-based monitoring tools are mentioned in the context of sediment quality assessment. Specific reference to the sediment Triad approach and TIE/EDA is made. In CIS guidance no. 27, a two tiered approach for the assessment of sediment quality is proposed.

Furthermore, the potential risk of cumulative effects from substances having the same mode of action should, according to CIS guidance no. 3, be taken into account in the pressures and impacts assessment, and according to the WFD, "other pollutants also need to be monitored if they are discharged in significant quantities in the river basin or subbasin." CIS guidance no. 7 states that "quantities that could compromise the

achievement of one of the Directive's objectives are clearly significant, and as examples, one might assume that a discharge that impacted a Protected Area, or caused exceedance of any national standard set under Annex V 1.2.6 of the Directive *or caused a biological or ecotoxicological effect* in a water body would be expected to be significant."



Look in: CIS guidance numbers 3, 7, 19, 25, 27

3.2 Objectives of effect-based monitoring in a WFD context

As with all other components of a monitoring programme, it is important to assess the suitability of different effect-based tools against the specified objectives of the whole programme (Chapman and Jackson, 1996). The suitability of any particular approach must be evaluated with respect to the cost and practicality of the method, and the ability of the method to provide information that can be translated into useful management information and to help achieve the monitoring programme objectives.

The focus of this chapter is to describe, in general terms, how the different categories of effect-based tools included in this report (see chapter 1.2. and chapters 4 and 5), can help Member States to implement in a more pragmatic way the objectives of the WFD.

As was already evident from previous CIS guidance documents, it is possible to identify several objectives for the use of effect-based tools in a WFD context, and a few of them are summarised below, together with suggested approaches.

- As screening tools, as part of the pressures and impacts assessment to aid in the prioritisation of water bodies to study further.
- To establish early warning systems, to prioritise further studies in areas that are not concluded to be at risk because they are located far from known local sources.
- To take the effects from mixtures of pollutants or not analysed chemicals into account (e.g. to support investigative monitoring where causes of a decline of specific species are unknown).
- To provide additional support in water and sediment quality assessment, though not as a replacement for conventional chemical and ecological monitoring under the WFD.

Effect-based tools are especially suitable as part of investigative monitoring programmes, for which the regulatory requirements are less formally determined. Each of these objectives and possible general approaches using the major categories of tools

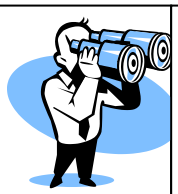
(summarised in chapter 1.2.) are described below. However, as with any investigative monitoring, the optimum set of tools to use varies on a case-to-case basis. The optimal approach will frequently involve several tools (including chemical analysis), as illustrated by several of the case studies found in the Annex (section 1) to this report. To optimise cost effectiveness, it is often wise to make use of the same samples for both chemical and effect-based analyses.

3.2.1 Prioritisation of water bodies for further monitoring

As mentioned earlier, the analysis of pressures and assessment of impacts is an important first step under the WFD in identifying operational monitoring needs. During this analysis, a very large number of pressures should be taken into account, including both point and diffuse sources. The analysis should result in conclusions about which water bodies are at risk of failing their specified objectives. As part of the pressures and impacts analysis, a list of relevant pollutants (including both priority substances and specific pollutants) needs to be identified.

To monitor primarily those substances that are emitted locally and make estimates based on emission data can become very difficult in practice. Even for large point sources, there is limited information on what hazardous substances (priority and specific) are actually emitted[22]. In addition there may be large numbers of known or potentially contaminated sites from historic activities, as well as several other local, regional and global diffuse sources that are less characterised but could release a complex mixture of substances and actually often can be expected to have substantial impact on the overall situation[23].

Emitted substances may also be more or less transformed into other substances. Thus, there is frequently a rather complex and variable mixture to take into account. The potential risk of cumulative effects from substances having the same mode of action should, according to CIS guidance no. 3, be taken into account[24]. There will be a very high level of uncertainty involved in the assessment of risks related to cumulative exposure, because of the limited chemical /modelled/ data for mixtures of partly unknown composition and with limited knowledge about the mode of action of the emitted substances.



Look in:

CIS guidance no. 3. Analysis of Pressures and Impacts

Table 3.9. The generic approach to the identification of specific pollutants

[22] Emission limit values are seldom developed for priority substances, although the recently implemented Industrial Emission Directive (2010/75/EC) does refer also to emissions of priority substances. However, several other point sources, most notably the sewage treatment plants, are not regulated by the IED.
[23] See e.g. conclusions made in the COHIBA project on the eleven substances specifically selected within the BSAP; http://www.cohiba-project.net/home/en_GB/home/
[24] See e.g. page 40 in CIS guidance no. 3

As part of the impact assessment, also monitoring data can be taken into account. By adding effect-based tools that respond to substances having the mode of actions specified in Annex VIII (in particular point 4[25] such as mutagenicity, impaired reproduction and endocrine functions), a large number of substances can be covered with only a few analyses and also taking cumulative effects into account. This is of particular importance in complex situations with many potential sources and insufficiently characterised emissions, and where the number of specific substances to consider is very large.

To cover a large number of potential substances having relevant modes of action, batteries of bioassays and/or biomarkers are suitable, combined with an initial screening of relevant priority substances. Time integrated information can be retrieved from biomarkers, sediment bioassays and/or *in vitro* batteries performed with passive sampling of water. This is of great value in situations where there are both continuous inputs via waste water discharges and highly fluctuating concentrations, such as seasonal runoff from plant protection products or other diffuse sources. *In vitro* assays and specific biomarkers can, combined with data on preselected priority substances, also provide information about the /type of/ variables to consider first in the selected water bodies.

Relevant case studies that illustrate some of the possibilities to use effect-based tools in this context are e.g.:

- "Endocrine disruptors in the Irish Aquatic Environment", on the use of *in vitro* and exposure biomarker tools to exclude that estrogenic substances are of concern to wild fish populations.

- "Laxsjön – investigating sediment contamination, using chemical and *in vitro* bioassay approach", on the use of combined chemical and *in vitro* bioassay battery to identify most impacted areas and relevant substances to study further.

- "Monitoring concentrated surface water with *in vivo* bioassays in the Netherlands", on the use of in vivo bioassay batteries combined with chemical analyses to identify toxic pressure.

- "The risk of chronic impact of pollution on the Bilinia river", on the use of *in vivo* and *in vitro* bioassays combined with chemical analyses to identify locations at risk.

- "Evaluation of Aquatic Environmental Oestrogens with Passive Sampling – EPSA", on the combined use of passive sampling and bioassays.

Refer also to

[25] Substances and preparations, or the breakdown products of such, which have been proved to possess: carcinogenic or mutagenic properties or properties which may affect steroidogenic, thyroid, reproduction or other endocrine-related functions in or via the aquatic environment.

- Chapter 4, on e.g. biomarker and *in vitro* batteries, specific biomarkers, sediment bioassays.

3.2.2 Early warning

It is of major importance to discover effects related to chemical substances before significant effects on population level occur, because damage at the population and ecosystem level can take a long time to repair. For example, at certain trophic levels, recolonisation can take much longer than the time frames (6 year management cycles) considered in the WFD. The ecological tools/indices in the WFD are not predictive, whereas effect-based tools, by detecting effects caused by chemical substances at an earlier stage, can help to serve a predictive role as regards higher-level effects.

Because of its focus being on local/regional pressures, the analysis of pressures and impacts will not primarily identify impacts in more remote areas. However, specific pollutants can be relevant also at large geographical scales and should be considered in surveillance monitoring programmes.

If associated with chemical stressors, alterations at the molecular and cellular levels can provide a sensitive indication of early changes that often represent the first warning signals of environmental disturbance. A multi-biomarker approach can for example be used to indicate risks of effects at population and community levels in the longer term. They can therefore be of value as a complementary tool to chemical analysis, to detect changes and identify hazards related to unknown toxic substances on a large geographical scale.

Several sensitive biomarker batteries have been used on a regular basis as early warning systems for several decades, especially in marine environments. Many of these are possible to apply also in freshwater environments. Experience can be gained from the use of such batteries in remote areas. A particular challenge however, is to identify causes and suitable control measures if significant effects are observed, since samples are from areas far from known local sources. Any warning signals would need to trigger investigative monitoring projects, such as effects directed analyses and regional studies upstream.

Relevant case studies that illustrate some of the possibilities to use effect-based tools in this context are:

- “Swedish national monitoring programme of fish health”, on biomarker batteries in fish, studying long term trends

Refer also to

- Chapter 4 on biomarker and bioassay batteries
- Chapter 6 on EDA

3.2.5. Effects from mixtures of pollutants and not analysed chemicals.

The Communication of the Commission of 31st of May 2012 states that in relation to the effects on wild species and ecosystems, the situation is less clear and the possibility of combination/mixture effects should be considered both in the case of independently acting chemicals as well as for chemicals with similar modes of action. Methodologies for the identification of chemical mixtures of potential concern as well as for the assessment of chemical mixtures are available. However, there are extensive knowledge and data gaps (mainly related to the mode of action and exposure data) that limit the extent to which mixtures can be properly assessed. The Commission will develop, by June 2014, and taking account of the opinion of the Scientific Committees, technical guidelines to promote a consistent approach to the assessment of priority mixtures across the different pieces of EU legislation. Effect-based tools integrate the effects that derive from pollutants (also those not analysed chemically) in the environment. In particular the presence of pollutants with the same mode of action (e.g. mutagenicity), although each at low-level concentrations, can be detected through the use of specific biomarkers for example.

Relevant case studies that illustrate some of the possibilities to use effect-based tools in this context are:

- “Monitoring concentrated surface water with in vivo bioassays in the Netherlands” on the study of toxic pressures using in vivo bioassays.

- “Endocrine Disruptors in the Irish Aquatic Environment” on the study of endocrine effects.

- “Swedish national monitoring programme of fish health” on the study of fish health and indications of effects from combined exposures measured using biomarkers.

- “Laxsjön – investigating sediment contamination, using chemical and *in vitro* bioassay approach” on the study of toxic pressures from combined exposures to substance having the same mode of action.

- “Deployment of a multi-biomarker approach to identify the origin of wild fish abnormalities reported in a French stream receiving urban and industrial effluents” on identification of causes to observed effects in fish.

Refer also to:

- Chapter 4 and Annex (section 7: “Biomarkers and *in vitro* assay related to certain modes of action”).

- Chapter 5 on ecological indicators (PICT and SPEAR).

3.2.6 Water and sediment quality assessment

3.2.6.1. Water quality assessment – chemical status

Chemical status assessment should, as was previously described, be based on chemical monitoring data regarding the list of priority substances. There are therefore limited possibilties (or reasons) to use effect-based tools in this context. However, chemical analysis of these substances is not always straightforward. For some substances the Limits of Quantification (LOQ) are not sufficient for compliance checking (cf requirements of 2009/90/EC), at least not if using routine analyses. While such tools are being developed and standardised, certain effect-based tools can provide a cost-effective screening-level approach to investigate water quality. The use of sufficiently sensitive chemical tools is required to preclude the presence of priority substances (and other substances listed in 2008/105/EC) at levels exceeding their corresponding EQS values. The potential use of certain *in vitro* assays[26] and sensitive exposure biomarkers[27] related to some of the priority substances is further discussed in chapter 4.

Relevant case studies that illustrate possibilities for using effect-based tools in this context are:

- "Endocrine Disruptors in the Irish Aquatic Environment", on the use of *in vitro* and exposure biomarker tools to exclude that estrogenic substances are of concern to wild fish populations.
- "Laxsjön – investigating sediment contamination, using chemical and *in vitro* bioassay approach", on the use of DR CALUX vs chemical analysis

Refer also to

- Chapter 4 (4.2.3.; 4.2.4.; 4.4.3.), on *in vitro* assays and exposure biomarkers related to estrogenic substances, as well as the dioxin reporter gene assay

3.2.6.2. Sediment quality assessment**-chemical status**

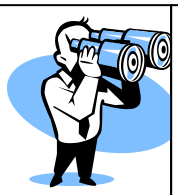
For some substances and types of water bodies (particularly coastal and lake), sediment can be the most suitable compartment to monitor, for practical, scientific and technical reasons (cf CIS guidance no. 25). Member States can check chemical and ecological status on the basis of national sediment EQS values. National sediment-EQS values should according to CIS guidance no. 27 be based only on the risks to benthic organisms (and not the other receptors such as secondary poisoning and human health). In practice,

[26] In particular oestrogen receptor transactivation assays (e.g. YES, ER CALUX) and dioxin reporter gene assay (e.g. DR CALUX) are proposed, related to the substances E-2, EE-2 and TCCDeq.

[27] Vitellogenin induction in male fish is e.g. considered to be highly sensitive early warning biomarker for exposure to oestrogenic substances; related to the substances E-2 and EE-2.

this means that they would frequently have to be based on effect levels in pelagic organisms but recalculated into sediment, because of a major lack in sediment toxicity data. The bioavailability of accumulated substances in sediment is not known and sediment- EQS values calculated using equilibrium partitioning theory might result in uncertainties. CIS guidance no. 27 therefore suggests a tiered assessment for such circumstances, by which a worst-case approach is used as a first-tier assessment, whereas bioavailability and/or effects are investigated in the second tier before enforcing potentially costly actions such as remediation. The Italian Ministerial Decree 260/2010 has already adopted this approach to evaluate sediment contamination in compliance with the CIS guidance no. 27.

Appropriate effect-based tools in such a second-tier assessment to assess bioavailability could include *in vivo* bioassays on either whole sediment or sediment pore water. Specific biomarkers could also indicate whether substances mainly found in sediment are bioavailable to benthic organisms. As for the current priority substances, there is in particular one specific biomarker, imposex in gastropods, that can be of high practical value to confirm TBT concentrations above EQS.



Look in:

CIS guidance document No 27. Technical Guidance for deriving Environmental Quality Standards

Chapter 5. Standards to protect benthic (sediment dwelling) species

Relevant case study that illustrates some of the possibilities to use effect-based tools in this context is:

- "Monitoring imposex on water body level", on the use of a specific effect biomarker to make conclusions about ecologically relevant effects from organic tin compounds analysed in sediments

Refer also to :

- Chapter 4 on specific biomarkers and sediment in vivo bioassays

3.2.6.3. Water and sediment quality assessment – Ecological Status

Because ecological status assessment should be based on already specified criteria (see chapter 2.3.) and need to fulfil certain quality requirements, any other tools can at the moment primarily be used as additional support.

As stated previously, ecological status should include an assessment related to the biological quality indicators. A few promising community-level tools, with a clear link to the effects from hazardous substances, are described in chapter 5. These ecological indicators show potential for future use in assessing ecological status.

Because a risk-based approach is promoted in ecological-status classification (requiring also the analysis of chemical variables), it can be assumed, at least from a scientific perspective, that bioassays and/or biomarkers (chapter 4) could be suitable to use to assess water quality, if the approach chosen is able to predict the risks of population and community-level relevant effects. This can be achieved by selecting ecolologically relevant variables and/or make an integrated assessment based on several variables.

In general, a biomarker approach based on "effect" and "general" biomarkers would have a high capacity to be used in this context, because of the highly integrated and ecologically relevant information obtained. ICES has also recently identified a suite of tools that are recommended in an integrated monitoring and assessment approach to assess environmental quality (Davies & Vethaak 2012). A battery of biomarkers and assays, and an integrated assessment approach is considered most suitable for this purpose (see chapter 4).

It can also be anticipated, at least from a scientific point of view, that *in vitro* bioassays that monitor WFD relevant modes of action (represented by the substances specified in Annex VIII), could specifically be used at least in a screening approach to assess risks of effects from these type of substances.

Relevant case studies that illustrate some of the possibilities to use effect-based tools in this context are:

- "Endocrine Disruptors in the Irish Aquatic Environment", on the use of *in vitro* and exposure biomarker tools to exclude that oestrogenic substances are of concern to wild fish populations.

- "Monitoring imposex on water body level", on the use of a specific effect biomarker to make conclusions about ecologically relevant effects from organic tin compounds

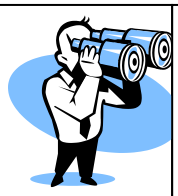
Refer also to :

- Chapter 4 on different types of biomarkers and the evaluation of biomarker batteries, and on the different modes of action possible to study using biomarkers and/or *in vitro* assays; integrated assessments

- Chapter 5 on ecological indicators

3.3 Effect-based tools within the MSFD

According to the COM decision 2010:477 a cause/effect relationship needs to be established for the monitoring tools to consider as 8.2.1. indicators.



Look at COM decision 2010:477 related to indicators to be developed:
"Levels of pollution effects on the ecosystem components concerned, having regard to the selected biological processes and taxonomic groups where a cause/effect relationship has been established and needs to be monitored (8.2.1)"

So far, no guidance documents have been established and the COM decision is currently under review. However, a valuable background document in this context is the so-called Task Group 8 report (Law et al 2010). See also chapter 2.1. on the ongoing work within the Regional Seas Conventions.

There are several monitoring tools for which there are established cause-effect relationships (see e.g. chapter 4, table 4.1. and 4.2.). To mention some examples, imposex biomarkers such as VSDI (respond to TBT), the Micronucleus assay (genotoxic compounds, such as certain PAHs[28]), VTG in male fish (oestrogenic compounds, such as EE2), EROD (PAHs and dioxin-like compounds), DNA adducts (PAHs) etc. There are also other tools, not described in this report but that are probably primarily of interest in a MSFD rather than WFD context, such as marine bird egg-shell thinning (responds to organochlorine pesticides such as DDT). Several tools are described in guidance documents developed by for example OSPAR (see e.g. JAMP guidelines for contaminant specific biological effects) and the Task Group 8 report.

However, several of the mentioned biomarkers (such as VTG and EROD) are traditionally primarily considered to be *exposure* biomarkers, rather than *effect* biomarkers, because a response does not necessarily reflect the onset of adverse health effects, at least not in a short-term perspective, see chapter 4.4. Considering that 8.2.1. should focus on effects of contaminants, the most appropriate way to include such biomarkers to assess GES would probably be as part of a weight-of-evidence or stepwise approach, together with other biomarkers.

Effect biomarkers, such as lysosomal stability and stress proteins, can respond to a wide number of stress factors, including also other factors than hazardous substances. However, they are frequently used to investigate effects from hazardous substances, see e.g. OSPAR JAMP guidelines for general biological effects monitoring. Nevertheless, such tools that are monitoring effects that are of high ecological relevance would be very useful in identifying areas of concern (low status). If responding to a spectrum of compounds and

[28] OECD standards are available and were developed for chemicals testing. Although the in vivo standard is developed for human erythrocytes, the method is also applicable to e.g. fish erythrocytes.

possibly also other types of stressors, conclusions as to whether the effects should primarily be considered to be related to hazardous substances would probably need an integrated (weight-of-evidence) approach and further investigations.

Only a very limited number of biomarkers could be considered single-substance specific; primarily ALA-D (Pb) and imposex (TBT), but not for example MT (responds to several metals). Not even for these very specific biomarkers however, can the absence of impact from other factors be ruled out. However, these tools provide valuable support in the evaluation of chemical data for these particular compounds.

For the evaluation phase, assessment criteria that can be used to establish GES are needed. Assessment criteria for individual marine effect-based monitoring tools have primarily been developed jointly by ICES and OSPAR, and are included in the Annex (section 8). Full details on how such background assessment criteria (BAC) and environmental assessment criteria (EAC) were derived can be found in the SGIMC 2010 and SGIMC 2011 and WKIMON 2009 reports and in the OSPAR Background Documents for individual biological effects methods. A compilation of the Background Documents and Assessment criteria is published in Davies & Vethaak 2012. Updates and amendments to the BACs and EACs can be found in the ICES WGBEC reports from 2012 and 2013 (ICES 2012; 2013). Values are reviewed annually and can be revised as new data becomes available.

In particular for persistent substances and substances considered under D9, the monitoring matrix of choice is frequently biota, and more specifically fish (edible tissues) and seafood. In such a context, also adding biomarker analyses of the same individuals or at least population, can offer a cost-effective, integrated monitoring approach (limit additional costs for sampling) and aid in the evaluation of data. ICES recently published a report on Integrated marine environmental monitoring of chemicals and their effects (Davies & Veethak 2012) that provides additional useful information and background information on the suggested different biomarker and bioassay components . As pointed out above, integrated or stepwise assessments are frequently also the most appropriate way to evaluate at least certain types of effect-based monitoring data, both to assess environmental quality and to identify the cause for the observed effects, in order to support the implementation of measures.

Relevant case studies that illustrate some of the possibilities to use effect-based tools in this context are:

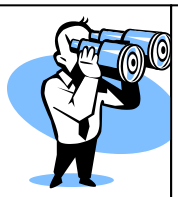
- “Swedish national monitoring programme of fish health” on the study and integrated assessment of biomarker data to assess fish health
- “Monitoring imposex on water body level”, on the use of a specific effect biomarker to make conclusions about ecologically relevant effects

Refer also to :

- Chapter 4 on biomarkers and the evaluation of data, including stepwise and integrated assessments
- Chapter 6 on EDA, to identify substances causing the effects

4 BIOMARKERS AND BIOASSAYS – TECHNICAL AND SCIENTIFIC ASPECTS

In the WFD Annex VIII there is an indicative list of main pollutants and point 4 and 5 in this list refer to substances with certain type of effects (mode of actions). Point 4 refers to substances which have been proved to possess carcinogenic or mutagenic properties, or properties which may affect steroidogenic, thyroid, reproduction or other endocrine- related functions in or via the aquatic environment. Point 5 refers to persistent hydrocarbons and persistent and bioaccumulable organic toxic substances.



Look at Annex VIII to the WFD

Point 4 and 5

In deriving environmental quality standards for pollutants listed in Annex VIII to the WFD, the provisions in WFD Annex V, 1.2.6. should be followed. CIS no. 27 provides further guidance on how to elaborate such standards. In particular, A1.3.3.14 of that guidance lists some relevant test endpoints (variables/effects studied) for toxicity tests to use in EQS derivation (after the addition of assessment/safety factors), because consequences at the population level of the test species are anticipated:

- growth (weight, length, growth rate, biomass)
- number (cells, population)
- mortality
- immobilisation
- reproduction
- hatching (rate, time, percentage)
- sex ratio
- development (egg, embryo, life stage)
- malformations (teratogenicity)
- proliferation (cells)
- filtration rate
- carbon uptake (algae)
- reburial (of e.g. certain crustacean species)

Several effect-based monitoring tools can identify samples that contain substances that possess such “WFD relevant modes of action” (e.g. mutagenic and endocrine disrupting properties) and type of effects.

The categorization in this report and chapter, into biomarkers, *in vitro* and *in vivo* bioassays is related to type of monitoring approach rather than the endpoint measured)[29]. Bioassays provide information about effects on subcellular levels or cells (*in vitro*) or living organisms (*in vivo*) when exposed to samples collected from the environment, generally water and sediment. Biomarkers (and ecological indicators, chapter 5) on the other hand, can be analysed in organisms from the field, integrating both exposure ("sampling") and effect and responding to those substances that are bioavailable in the environment.

The ways effect-based tools are used at national level today and how they can become valuable in the WFD monitoring context has been described briefly in chapters 2 and 3. To a large extent, such monitoring approaches would today primarily involve biomarkers and bioassays, tools that measure effects on individuals and at suborganism level. Several such tools are already used on a routine basis and this chapter is focused on describing some of the technical and scientific aspects. More than twenty individual biomarkers and eight *in vitro* bioassay tools are also described in some detail in fact sheets found in the Annex (section 6). The fact sheets include references to methods used, costs, type of environment that can be investigated (marine-limnic, river-lake), QA procedures, specificity, other influencing factors, as well as ecological relevance.

4.1 Bioassays

Bioassays measure the toxicity of environmental samples, generally under defined laboratory conditions, using a common procedure to measure toxicological endpoints at organism level or below [Piva et al., 2011 and references therein]. They are preferably applied in a battery, using a large number of test species from several taxa, and across the main ecological or trophic positions (i.e. from bacteria to fish, and from decomposers to final consumers).

In this report, bioassays are divided into two types:

- *in vitro* assays (exposed cell lines)
- *in vivo* whole organism bioassays (although endpoints can include *in vitro* studies)

4.1.1 Sample pretreatment

Both *in vitro* and in vivo bioassays generally require environmental samples to be collected[30], most frequently from water and/or sediment but also samples originating from biota (e.g. bile from fish or lipid extracted from mussels). Therefore, as with chemical analyses, the sampling frequency and pre-treatment of environmental samples are

[29] To illustrate, if EROD induction is analysed on liver samples from wild caught fish, it is in this report considered a biomarker. The study of EROD induction in liver cells cultured and exposed to samples collected in the field (or effluents), EROD is considered an *in vitro* bioassay. Finally, if studying EROD induction on liver samples from fish kept in the laboratory and exposed to field collected samples, such analyses are considered to be *in vitro* analyses but within the context of an *in vivo* bioassay. The borders are not always clear. Analysing EROD induction in liver samples from caged fish in the field, would still be called a biomarker in this report but the study is performed as an in situ bioassay. Other combintations exist, such as the monitoring of organisms in the laboratory but exposing them to a flow-through system based on continuous field sampling (e.g. the commercial toximeters that exist for fish, daphnia and algae and that are used as alarm systems to protect drinking water, cf capter 2 on national use).

[30] However, *in vivo* bioassays are also possible to perform in the field, then generally referred to as *in situ* testing.

important aspects to consider because any pretreatment can cause an alteration of the chemical composition of a sample. There are also general aspects associated with the sampling process itself, for example, the frequency of sampling and quality assurance surrounding sampling. This topic has been dealt with elsewhere, and in most cases the aspects to consider are the same, regardless of whether the samples are analysed chemically or toxicologically. This will therefore not be described in detail in this report. For example, Madrid and Zayas (2007) discuss water strategies in relation to the WFD and list a number of ISO guidelines for sampling (e.g. ISO 5667-6:2005 for sampling rivers and streams; see also CIS guidance documents nos. 7 and 25).

In particular for water samples though, there are some critical issues to be considered specifically when analysing using bioassays, and in particular *in vivo* bioassays because it is frequently necessary to perform some kind of pretreatment of the sample, unless acute effects can be expected to occur. Acute effects are less evident in European surface water samples, although they can be detected in contaminated sites or where there is a relevant unpredictable accident that requires investigative monitoring. These aspects are therefore described in more detail in the Annex (section 3).

While the quality of surface waters in Europe has significantly improved during the past years, highly contaminated sediments still create a considerable threat to the quality of several European catchment areas (Brinkmann et al. 2010). Sediment contains an integrated quantity of contaminants, and concentrations of contaminants in sediments will show much less temporal variability over time when compared to the overlying water phase. Short-term *in vivo* bioassays can be useful in a monitoring context also to detect possible effects on benthic organisms. As with water samples, native (whole) sediment samples can be tested directly with *in vivo* and *in vitro* bioassays. In whole sediment bioassay testing, the sample is exposed to minimum disturbance, besides adding standard test water, in order to simulate conditions in situ, and minimize potential impact on bioavailability of chemicals present in the sample (due to e.g. “aging”[31]), but still perform the test under standardized conditions. Alternatively, either pore water, elutriates or sediment extracts can be used in the bioassays (Ahlf et al. 2002). The choice between sediment pretreatment or not, is primarily related to the purpose of the testing, but also the choice of test organism. There are for example difficulties to observe behavioural changes in small test organisms if studied in whole sediment assays. Within the last years several weight-of-evidence studies using mechanism-specific bioassays have been employed in order to link the toxicity of sediment-bound pollutants and the situation *in situ* in the field (eg, Keiter et al, 2006, Böttcher et al. 2010).

4.1.2 Battery of bioassays

For a broad scope, the battery of ecotoxicological tests should have a sensitive and an overall discriminatory power responding to as many forms of contamination as possible.The most suitable approach is generally based on the choice of an adequate battery of tests and choice of species which should take into account different aspects: sensitivity, specificity, availability of organisms (for *in vivo* bioassays), the variability of the method, cost/effectiveness, ethics, as well as standardization and intercalibration of the methods. Although *in vitro* and *in vivo* assays are described separately below, integrated

[31] A prolonged contact time between e.g. PAHs and organic carbon in sediment may cause stronger binding (Landrum et al 1992)

monitoring approaches using both types of assays are also common and they frequently include also chemical analyses being performed on the same samples, see e.g. case studies “The risk of chronic impact of pollution on the Bílina River”; “Monitoring concentrated surface water with *in vivo* bioassays” in Annex (section 1) and Annex section 2.

4.1.3 Validation

Standardisation and intercalibration aspects are of particular importance if monitoring results are to be used in a regulatory context, also emphasised by the 2009/90/EC in the WFD context. Internationally, the OECD (Organisation for Economic Co-operation and development) and ISO (International Organisation of Standardisation) are the most important bodies for development, validation and standardization of analytical as well as effect-based test methods. Whereas the Test Guidelines Programme, within the Environmental Directorate of the OECD is focused on test methods for single substance testing (“toxicity tests”), the Technical Committee (TC) 147 “Water Quality” of ISO is dedicated to the environmental aspects of water quality control. Another important body of validation and standardization of bioassays/toxicity tests is the US EPA. In general, the protocols for single substance tests can frequently be adapted to work also for complex environmental samples. However, environmental samples usually have much lower concentrations of toxic substances (see above) than the concentration ranges generally used in toxicity tests within chemicals testing[32]. The Annex (section 4) to this report includes a table that lists several *in vivo* bioassays/toxicity tests but also some *in vitro* assays and biomarkers for which there are now standards available.

Further standardisation of effect-based tools such as *in vitro* bioassays for regulatory applications and use for surface water monitoring is needed (Kase et al., 2009). For investigative purposes such as screening however, non standardised methods could still be very valuable, and are sometimes the only option also within chemical analysis. A common validation framework that can also cover tools that are less established is therefore valuable to increase comparability of such data from different regions. The NORMAN network has for example developed a common framework for the validation of chemical and biological monitoring methods, to investigate occurrence and effects of emerging pollutants[33]. Three types of validation aspects are distinguished in this context: within-laboratory validation (research level), basic external validation (transferability at expert level) and, the most complex level: inter-laboratory validation on routine level. The NORMAN protocol is under negotiation at CEN level (CEN TC 230 - Water Analysis). For more details on the procedures related to standardisation in general but aspects to consider for effect-based tools in particular, see Annex section 5.

[32] Within chemicals testing, the highest test concentrations are only limited by the solubility of the tested substance.

[33] http://www.norman-network.net/index_php.php?module=public/qa_qc/validation&menu2=public/qa_qc/qa_qc

4.1.4 Working with measures – how to identify the cause

Although bioassays do not immediately provide information on the underlying substances causing an effect, bioassays are performed on samples taken from the environment, and in this sense, the same types of back tracking investigation can be performed as when using a chemical approach. By sampling in a gradient, if primarily local sources are suspected, the same difficulties are involved as with a chemical approach. By repeated testing/analysis of samples from gradients (and possibly effluents etc) local sources can be identified. In some cases though, it is necessary to identify the substance/s that are causing the effects. Although more laborious, a potential approach to obtain at least a rough estimate and sometimes the exact identities of the causative agents is to use a combination of chemical and toxicological tools (see chapter 6 on EDA).

4.2 In vitro bioassays – general technical considerations

Many of the *in vitro* bioassay protocols originate from human toxicity screening tests developed for chemical regulation purposes. The use of *in vitro* assays is increasing for ethical reasons in compliance with the laws about animal experimentation. They measure effects on lower organisational levels (such as receptor activation and DNA damage). However, instead of investigating cells from tissues of organisms that were exposed in the field or from caged whole organisms (as is the case with biomarkers), the effects are studied on cell lines after exposure to environmental samples. An advantage is that *in vitro* bioassays can frequently be performed on any[34] matrix (such as /extracts of/ surface water, sediment and pore water, biological tissues, passive samplers and effluents).

Additional advantages are that only small amounts of sample (grams) are generally needed and the exposure time is generally short compared to the time needed in an *in vivo* assay to detect a response from substances having the same mode of action. In most cases, *in vitro* assays are considered sensitive, because they measure effects on a low organisational level. Many (but not all) *in vitro* assays are suitable for screening and high throughput and automated applications and can be added to the analytical package at comparatively low costs (especially if taking into account the number of substances they respond to).

For some *in vitro* bioassays the results are expressed in chemical equivalents, referring to the response from the sample compared to the response from the reference chemical (positive control). However, before comparing such results to the criteria developed for a single chemical or a specified number of chemicals, it is necessary to consider that the assay response is related to several substances combined having the same mode of action, e.g. via receptor activation. Generally for similar acting agonistic substances, the biological signal is higher than the chemical single substance based signal which makes the in vitro assays suitable as screening tools. As integrative detection tools, in vitro assays are also able to quantify and distinguish agonistic and antagonistic effects. For example by adding a reference substance to the sample it is possible to test and identify

[34] Although *in vitro* bioassays can be used on any matrix /extract/, some are more suited for the assessment of certain matrixes than others, in part because they were so far only validated for certain uses, but also because relevant substances that elicit certain types of responses primarily are found in certain compartments.

receptor inhibiting antagonistic effects in environmental samples, due to a less potent receptor binding in comparison to the reference substance testing.

For simultaneously acting antagonists the biological signal response can be reduced, but the effect can be identified in the same testing step, e.g. on the same microtiter plate in parallel. Furthermore, the possibility to compare the results directly with the corresponding EQS other than for screening purposes varies depending on the basis for the EQS value.

Some drawbacks with *in vitro* tools are that, as opposed to *in vivo* bioassays and biomarkers, the systems studied are highly simplified compared to the complexity of whole organisms. Thus the interaction between different receptors, cells and organs is not studied. Such aspects can have important implications in interpreting the results. The substance 17-alpha-ethinylestradiol (EE2) has for example. a slightly higher potency *in vitro* than the natural hormone 17-beta-estradiol (E2), but *in vivo* (using biomarkers) it is considered 10-25 times more potent (further discussed in 4.2.3.). As in chemical analysis (and as opposed to biomarkers measured on field exposed organisms providing more integrated responses), one can only detect effects from substances that are present in the sample. Bioavailablity is also difficult to assess unless analysing biological tissues[35].

4.2.1 Working with measures – how to identify the cause

In vitro assays are suitable and sometimes necessary if wishing to make follow up studies related to effects observed using biomarkers (see Annex section 7). They can, in comparison to most biomarkers, easily be used to track local pollution sources by sampling water and sediment in a gradient but also effluents from suspected sources. *In vitro* assays are also very valuable in EDA/TIE approaches to identify toxic fractions and guide in identifying causative agents (see e.g. case study “Contaminated sediments in the River Elbe basin - EDA”).

4.2.2 *In vitro* bioassays available

Today, there are large numbers of *in vitro* bioassays available, in addition to those that are already standardised or being considered for standardisation. In principle, all *in vitro* analyses that can be performed as biomarkers, would likely also be possible to perform as *in vitro* bioassays, if cell lines are available. To cover all *in vitro* bioassays (and biomarkers) is out of the scope of this report. Furthermore, not all available in vitro assays can be considered easy to perform. Nevertheless, to mention a few common tools within monitoring, assays that have been initially selected for toxicity characterisation and EDA in the MODELKEY[36] project (Thomas 2006), recommended by COHIBA (2010) for whole effluent assessment, genotoxicity assays recommended by OSPAR (2002)[37] (both for

[35] On the other hand, in these cases, if measuring a hormonal response such as “endocrine disruption” care should be taken in the interpretation of *in vitro* assay data because of the possibilities to detect effects from endogenous hormones rather than xenobiotics. Furthermore, effects from substances that are available but do not accumulate, are difficult to detect by chemical analysis as well as *in vitro* assays of tissues (as opposed to biomarkers measured on field exposed organisms).

[36] http://www.modelkey.org/

[37] The OSPAR Commission (2002) recommends a test battery of bacterial assays (umu C or SOS chromo assay and Ames) and eukaryotic cells (micronucleus or Comet assay) for WEA.

whole effluent assessment and surface water monitoring), and assays that were nominated for aquatic monitoring purposes during a Swedish workshop[38] are included in table 4.1. It should be noted that the listing of certain *in vitro* bioassays in the table, does not implicate that they are necessarily recommended in a general sense.

Table 4.1. *In vitro* assays that were nominated for monitoring purposes during a Swedish Workshop (W), recommended for WEA assessments by COHIBA (C) or OSPAR (O), and initially selected for evaluation regarding high throughput screening and EDA purposes in the MODELKEY project (M[39]). The table also includes information about the type of compounds (mode of action) the assay responds to.

Name/s of assay	Workshop/ COHIBA/Modelkey	Mode of action/endpoint
AR CALUX (anti-)	W, **M**	Androgen receptor (activation or blocking)
DR CALUX	W, **M**	AH receptor binding
ER CALUX[40] (anti-)	W, **M**	/Alpha and beta/ oestrogen receptors
GR CALUX (anti-)	W	Glucocorticoid receptor
PAH CALUX	W, M[41]	AH receptor binding
PR CALUX	W	Progesterone receptor
Acetylcholinesterase inhibition assay	W	Inhibition of acetylcholinesterase activity
Carboxylesterase inhibition assay	W	Inhibition of carboxylesterase activity
Ames	W, **M**, O	Genotoxicity: Mutations[42]
umuC	W, **M**, C	SOS response to DNA damage[43]
TTR-binding	W, **M**	Competition with thyroid hormone for binding to TTR (transport protein)
TRb CALUX	W	Thyroid receptor beta
EROD	C	EROD induction
YES	C, **M**	ER receptor
YAS	C, **M**	AR receptor
P-53 accumulation	(M)[44]	Genotoxicity
Green screen	(M)[45]	Genotoxicity
RYA	**M**	ER receptor
ABC assay	**M**	Antibiotic activity

[38] « Effect based monitoring tools and assessment criteria". Göteborg, 25th–26th January 2011. National expertise, including primarily researchers but also regulators, consultants and representatives from commercial laboratories, was invited. Upon registration, participants were asked to "nominate" at least one monitoring tool that they thought should be considered for evaluation regarding monitoring purposes in the Swedish environment.More information on the outcome of the workshop can be found in Wernersson 2012. Swedish monitoring of hazardous substances in the aquatic environment - current vs required monitoring and potential developments. Länsstyrelsen Västra Götaland report no 2012:23.

[39] If "M" is typed in bold, the assay was considered to be useful for both water, sediment and tissue bioassays.

[40] There are actually two different types; ER and ERalpha, depending on the receptor activated/inhibited, see fact sheet

[41] Considered suitable for sediment and tissues

[42] Responds to reactivation of bacteria (Salmonella typhimurium) that can grow without histidine. Frequently used within WEA, german standard for this purpose. Microplate tests are available. TA 98 measures frame shift mutations; TA 100 base substitutions. Further strains, some genetically modified in order to express genes of the xenobiotic metabolism are available. External metabolic activation of chemicals by S9-mix.Refer to standard for more information

[43] The umuC assay measures the induction of the bacterial DNA reparation system (SOS) and is based on the reporter gene lac Z (beta galactosidase is formed). The assay is routinely used within WEA in Germany and there is also a German standard available (DIN 38415-4;1996). The test variant described by Grummt et al 2000 on surface water samples could also detect genotoxicity of surface water samples from four locations in the Elbe and Rhine. Refer to standard for more information

[44] Considered suitable for sediment assays only

[45] Considered suitable for water assays only

Name/s of assay	Workshop/ COHIBA/Modelkey	Mode of action/endpoint
Micronucleus test	C	Genotoxicity: Damage to chromosomes or mitotic apparatus
Vitellogenin induction test	C	Vitellogenin production
PPARy2 CALUX (anti-)	(W)[46]	Peroxisome proliferator activated receptors (PPARs)
Comet Assay	O	Genotoxicity: DNA damage monitored directly[47]

The usefulness of individual *in vitro* bioassays in a certain context frequently needs to be evaluated on a case-by-case basis, and some case studies are included in the Annex (section 1) (see e.g. "Laxsjön – investigating sediment contamination, using chemical and *in vitro* bioassay approach"; "The risk of chronic impact of pollution on the Bílina River"). Several literature reviews are also available, that include also other assays than the ones mentioned above. Behnisch et al (2001) reviewed different areas of applying *in vitro* bioassays in screening studies of dioxin and dioxinlike compounds. Lilja et al (2010) reviewed some of the available *in vitro* bioassays related to genetic effects and endocrine disruption regarding their usefulness in STP water monitoring.

The frequently used commercial CALUX (Chemical Activated Luciferase Gene Expression) panel is based on a reporter gene approach and the assays produce light when exposed to substances that induce certain pathways, such as Ah or oestrogen receptor (ER) binding. The molecule-receptor complex binds to specific DNA sequences (called "responsive elements"), triggering the expression of certain genes, in turn giving rise to the toxicological response. Also non commercial alternative cell lines are available, such as the T47D-Kbluc cell line.

The use of genotoxicity assays for environmental monitoring purposes (surface waters) was evaluated by Grummt et al 2000, and positive water samples were most frequently identified using the Comet assay although genotoxicity was also identified by new test variants of Ames and UmuC.

Yeast cell based assays, such as YES and YAS (recombinant) are being used more frequently, especially to screen effluent samples regarding oestrogenic and androgenic compounds. Kinnberg (2003) evaluated several *in vitro* assays, including the YES/YAS, ER CALUX[48] and E screen assays, for determination of oestrogenic activity in the environment. Leusch et al (2010) evaluated five oestrogenic assays regarding their usefulness in monitoring. A validation study of tools to determine oestrogens in sewage treatment effluents, including E-screen, were also performed by the NORMAN network (NORMAN 2008). The usefulness of a few selected *in vitro* assays in a WFD surface water monitoring context, to detect oestrogenic substances and Ah receptor binding substances respectively is described in more depth below.

[46] Not evaluated during the workshop but included in the discussion.
[47] By staining DNA from eucaryotic cells, exposed either *in vivo* or *in vitro* (permanent cell lines, frequently human hepatoma Hep GS) (OSPAR 2002).
[48] The initial ER CALUX assay, not alpha version which was developed at a later stage.

4.2.3 *In vitro* tools to detect the presence of endocrine disruptors in a WFD context

Some endocrine disruptive compounds (EDCs) influence the sexual function and differentiation in aquatic organisms, mainly driven by their oestrogenic or androgenic activity. A well-studied mode of action is oestrogenic receptor binding. In a recent statement, the European Commission recognized the importance of assessing the endocrine disrupting potential of individual chemicals as well as, where appropriate, the cumulative impact of identified combinations of substances on the endocrine system (European Commission 2011). Biomarkers that can be used to study endocrine disruption are further described in chapter 4.3.3. whereas this section is focused on possibilities for using *in vitro* bioassays on a screening level to detect areas at risk.

The oestrogenic substances 17-alpha-ethinylestradiol (EE2) and 17-beta-estradiol (E2) were listed as candidate priority substances with EQS values of 35 pg/l for EE2 and 0.4 ng/l for E2, respectively (COM 2011, 876). Under the new Directive 2013/39/EU these substances are to be included in the first EU-wide watchlist monitoring programme. Even though these substances have not been designated as priority substances in the 2013 Directive, they could be following a future review, and could be designated as RBSPs by individual Member States.

However, both EQS values are below the analytical limits of quantification (LOQ) of most routine chemical methods[49]. To overcome these current chemical detection problems, *in vitro* oestrogen-receptor transactivation assays can be used for the screening of oestrogenic activity, in order to identify water bodies at risk due to combined exposure to several oestrogenic substances, some of which could constitute potential RBSPs.

Three widely used oestrogen receptor transactivation assays have been suggested as suitable tools for monitoring oestrogenic activity in water samples (Kase et al., 2009; 2011, Hecker & Hollert, 2011; Kienle et al., 2011; 2012). These assays have also been compared in several studies (Murk et al., 2002; Leusch et al., 2008; Kase et al., 2009):

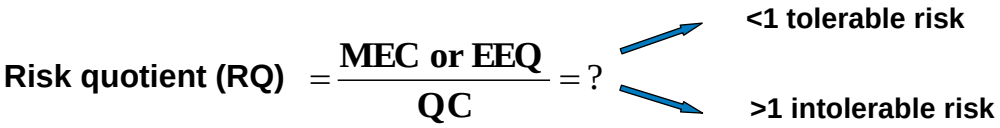
1. The **YES (Yeast Estrogen Screen) assay** (Routledge & Sumpter, 1996 adapted by Schultis & Metzger, 2004),
2. The commercial **ER-CALUX®** (Estrogen Receptor-mediated Chemically Activated Luciferace gene expression) (van der Linden et al., 2008), and
3. The non-commercial **T47D-Kbluc assay** (Wilson et al., 2004).

These bioanalytical methods have proven functionality in environmental samples and can be used for surface water assessment or to assess significant sources of potential endocrine disruptors such as municipial wastewater (Kienle et al., 2011) or sediments (Grund et al., 2011 ; see also the case study « Laxsjön – investigating sediment contamination, using chemical and *in vitro* bioassay approach").

These *in vitro* assays can be used to measure the overall receptor binding potential related to the combined potency of oestrogens, e.g. EE2, E2 or estrone (E1) and other substances in an environmental sample. The results are commonly expressed as E2-equivalents (« EEQs ») in bioanalytics and biomonitoring. E2 has an *in vitro* and *in vivo*

[49] The lowest detection limit reported by Loos 2012 was e.g. 0.1 ng/l for EE-2 and E-2, if using USEPA method 1698; in practice the LOQ that is possible to reach by regular laboratories is generally higher

potency between E1 (estrone) and EE2. In the report by Loos (2012), Kase et al therefore suggest that the results (MEC, Measured Environmental Concentration, or EEQ based on *in vitro* assay results) can be compared with the proposed EQS developed for E2, to determine related risk quotients (equation 1).



Equation 1: Calculation of risk quotients (RQ), MEC= Measured Environmental Concentration or equivalent concentration, e.g. EEQ; QC= Quality criteria (usually the AA-EQS)

Among the three assays, the **YES assay, in the DIN/ISO** standardisation programme, was generally found to be the least sensitive with an LOQ for E2 in the low ng/l range (see fact sheet). However, the advantages of the YES are its practicability and robustness also for waste water assessments. In such context, a prediction of potential anthropogenic oestrogenic impact on surface water can be made using the YES by dividing the EEQs by a corresponding dilution factor.

The **ER-CALUX®** and the non-commercial **T47D-Kbluc** are more sensitive than the YES. They reliably detect oestrogenic activity in surface water at sufficiently low concentrations (below ng/l) compared with the suggested EQS (see fact sheet).

All three *in vitro* assays can be performed in combination with solid phase extraction (SPE) and passive sampling, so lower LOQs are also possible, depending on the methods used. Different SPE-LOQs of **ER-CALUX®** of 20-40 pg/l are described (e.g. Puijker, 2007). The sensitivity of the **T47D-Kbluc** is expected to be in a similar range.

However, it must be stressed that these methods are integrative receptor binding assays which detect *all* oestrogen-like chemicals able to bind (agonistic) to the oestrogen receptor. Therefore, they can be considered suitable screening assays for both the overall oestrogenic potential and the single strongly binding substances such as EE2 and E2. As mentioned in chapter 4.2 the presence of antioestrogenic substances can be monitored in the same testing approach.

It should also be pointed out that because EE2 is significantly (about 10-25 times) more potent *in vivo* than E2, but only 2-3 times more, or equally potent in the *in vitro* assays mentioned, this should be taken into account if evaluating data in an absolute manner (comparison with EQS). To take this difference into account, an option could be to add additional safety factors before comparison (depending on the test used), if the presence of EE2 is likely to be of concern; or defining an appropriate trigger value based on the comparison of analytics and bioanalytics. Using a such stricter EEQ-value the potential presence of EE2 at EQS level can be screened in water bodies, where a significant exposure of waste water relevant micropollutants like EE2 is expected.

A combination of chemical measurements (of E2 and EE2) and *in vitro* bioassays (that also respond to other substances that bind to oestrogenic receptors) would be the most appropriate option to assess the quantitative risks from oestrogenic substances. However,

by using *in vitro* bioassays on an initial screening level, they are useful to lower the need for and frequency of analytical high end monitoring to those water bodies that are primarily considered to be at risk. Using *in vitro* bioassays could therefore reduce the high costs[50] of the few currently available analytical "high end" methods for the measurement of E2 and EE2. Samples that require chemical analytical confirmation for single compounds, can be further analysed with more sensitive (and costly) chromatographic analytical methods (based on LC- or GC-MS techniques) with LOQs below the recommended AA-EQS for E2 or EE2. Other known (and generally weaker) oestrogen receptor binding compounds, such as oestrone (E1), nonylphenols, bisphenol A and others should also be considered in such an analysis.

These tools have also successfully been used for the identification of unknown chemicals and the contribution of single compounds to the overall endocrine effectiveness when combined with the strategy of EDA (Hecker & Hollert, 2009; Higley et al., 2012), see chapter 6. The fractionation of samples is also a possible way to identify simultaneously acting substances having agonistic and antagonistic properties (oestrogenic and anti-oestrogenic effects), in a kind of mini EDA.

In addition to the monitoring for oestrogenic activities, androgenic activities can be monitored by specific androgen-receptor transactivation assays in parallel. The knowledge about androgenic and anti-androgenic receptor binding in the aquatic environment is currently limited, which is also the case for other receptor mediated activations (e.g. Kortenkamp et al., 2011). Therefore, additional monitoring with androgen receptor (AR) transactivation assays is preferably performed in parallel to address both ER- and AR-receptor mediated risks of endocrine disruptors with effect- based tools. It is known that other environmentally relevant water pollutants, e.g. triclosan, can increase oestrogenic activity via an inhibition of the androgen receptor (AR) (Rostowski et al., 2011). Therefore a simultaneous monitoring of ER and AR receptor activation and inhibition is preferred. Androgenic activities can be monitored by commercial **AR-CALUX®** systems, or the non-commercial **MDA-kb2 cell line** with an AR receptor, recommended by the US-EPA (Wilson et al., 2002; Blake et al., 2010; Hecker & Hollert, 2011). Similar to the EEQ approach androgenic hormone equivalents (AEQ) like testosterone, or dihydrotestosterone equivalents can be used as a positive control to calculate the AEQs .

At present no standardized instruments are in place for the detection of EE2 or E2 in routine monitoring of coastal and continental surface waters. An innovative monitoring strategy for both substances needs to be put forward, validated and implemented.The Federal Institute of Hydrology in Germany (BfG), the Federal Environment Agency (UBA) and the Swiss Centre for Applied Ecotoxicology invited experts from academia and authorities as well as representatives from CMEP[51] and DG JRC[52], to provide an expert opinion on the suitability of bioassays for the monitoring of EE2 and E2 in surface waters. The main conclusions are listed in the Annex section 2, and were presented at the 17th CIS WGE meeting[53].

[50] The bioanalytical cost range is between 60-200 Euro ; see fact sheets.
[51] "Chemical Monitoring and Emerging Pollutants"; sub-group under the Common Implementation Strategy for the Water Framework Directive
[52] Joint Research Centre
[53] 22nd April 2013 in Brussels

4.2.4 *In vitro* tools to detect the presence of Ah receptor binding substances in a WFD context

The direct measurement of the biological responses associated to the Ah receptor in samples provide, conversely to chemical analysis, a proper and rapid quantification of the overall potency of dioxin and dioxin-like compounds in the sample (Brack *et al.*, 2005; Brack *et al.*, 2007; Louiz *et al.*, 2008; Kinani *et al.*, 2010). Ah receptor binding assays respond to dioxins, dibenzofurans and planar PCBs, all new priority substances. However, they also respond to other persistent substances with the same mode of action, being of equal high levels of concern (such as brominated furans). Ah receptor binding assays are therefore suitable as screening tools, overestimating rather than underestimating the risks and indicating whether "non listed" substances should be considered.

The application of *in vitro* bioassays in screening studies of dioxin and dioxin-like compounds was for example reviewed by Behnisch et al (2001). One possible tool for measuring this activity is the commercial DR CALUX assay, or the non commercial alternative cell line MDA-kb2 with an Ah receptor, recommended by the US-EPA (Wilson et al. 2002, Blake et al 2010 and Hecker & Hollert 2011). Other methods based on a similar mechanism are e.g. PLHC 1-EROD, and RTL-W1 that have been used to assess dioxin-like contamination in environmental samples. The *in vitro* bioassays mentioned were used both for screening of sediments and biota (see e.g. Kinani et al, 2010; Traven et al., 1998; Hurst et al., 2004; Brack et al., 2005 and Thomas et al 2006), of particular concern in a WFD context, to predict risks in or via the aquatic environment.

Results are expressed as 2,3,7,8-TCDD TEQ and can therefore, in a screening context, be directly compared to the EQS (expressed in the same unit)[54]. Such an approach is already acceptable to estimate health risks from fish and other seafood according to current EU regulation (1881/2006), and the now established biota EQS is based on values developed in this context.

The costs are generally substantially lower for *in vitro* bioassays than chemical analyses of the regulated compounds (see e.g. fact sheet on DR CALUX). Therefore, there are generally economic reasons for using *in vitro* bioassays on a screening level.

In contrast to the case related to oestrogenic compounds described above, it is safer to make conclusions about the impact and risks from persistent dioxin-like substances based on *in vitro* bioassays rather than biomarker data (effects observed on e.g. cytochrome P450 1A system[55]). Indeed, CYP1A induction is due to activation of Ah receptor by dioxin-like substances and is described as a relevant biomarker of dioxin-like exposure (for review, see Whyte *et al.*, 2000). However, there are confounding and/or inhibiting factors that could influence CYP1A activity in a multi-contamination context, and activity is also induced by less persistent substances, such as certain PAHs that can also bind to the Ah receptor (see e.g. fact sheet on EROD). Furthermore, in a WFD context, because persistent dioxin-like compounds (dioxins, dibenzofurans and planar PCBs) are subject to biomagnification, the highest concern is related to biota at higher trophic levels, including humans, predatory birds and mammals. Hence, the use of fish biomarkers to assess the

[54] The joint effects of dioxins and dioxin like compounds are considered to act in an additive manner, and are therefore quantified on the basis of their relative potencies using the TEF approach.

[55] gene expression, protein or catalytic activity inductions

risks from dioxin-like compounds via the aquatic environment is not recommended in a WFD context.

As already mentioned, certain PAHs (such as benzo(a)pyrene), can also bind to the Ah receptor. In the traditional *in vitro* bioassays of Ah receptor binding activity, these are therefore removed from the sample before analysis, to make sure only persistent ("dioxin-like") substances are included. In recent years however, options to also discover Ah receptor binding PAHs by using the same test but with other pretreatment have evolved (see e.g. PAH CALUX fact sheet, as well as the case study "Laxsjön – investigating sediment contamination, using chemical and *in vitro* bioassay approach"). The assay is valuable in identifying samples having elevated levels of Ah receptor binding PAHs, but care should be taken before making conclusions about PAH concentrations. The assay most likely responds to the cumulative exposure to a large number of ("non parent") PAHs, i.e. other PAHs than those to consider within chemical status classifications. Nevertheless, the assay can be useful to locate water bodies at risk due to the presence of Ah receptor binding PAHs and biomarker studies could be a relevant second step to investigate impact on pelagic organisms from such compounds.

4.3 In vivo bioassays – general technical considerations

In vivo bioassays are tests in which whole living organisms (including bacteria) are exposed to environmental samples like surface water, sediment, waste water, dredge material or extracts from these samples. Tests are performed in the laboratory or, less frequently, in the field ("*in situ*" bioassays) (Rijkswaterstaat, 2005).

The "end point" is the type of effect that is measured in such a toxicological test, and some examples that are frequently used in this context are:

- Mortality
- Immobilization
- Effects on reproduction (i.e fertilization, hatching, embryo development)
- Effects on growth of individuals
- Effects on growth of populations
- Metabolic or physiological changes
- Behavioural changes
- Bioluminescence
- Molecular/Biochemical responses

In general, *in vivo* bioassays are broad spectrum assays, e.g an *in vivo* bioassay reacts on a variety of substances and on different types of toxicity. Nevertheless, it is important that the evaluation of toxic effects observed is based on the response observed in several species, because they can exhibit intrinsic differences in terms of sensitivity to various chemicals, also depending on the endpoint measured in the test [Ahlf & Heise, 2005]. Both short and long term *in vivo* bioassays should preferably be carried out and at least three species of organisms belonging to different taxonomic groups and trophic levels (primary producer, decomposer/saprophytic, detritivore/filter feeder, consumer). The battery of ecotoxicological tests should have sufficient sensitivity and an overall discriminatory power responding to as many forms of pollution as possible.

Acute toxicity means adverse effects that occur in a short time (not exceeding one third of the average time between birth and sexual maturity) while for chronic toxicity effects are measured after a period longer than 50% of the organism life time [Perin, 2004]. According to these definitions, it is not possible to establish a time limit, e.g. 24, 48 or 96 h, to distinguish between acute, subacute or chronic assays. Bacteria, algae, invertebrates or other model species may have a very different average life time [ECETOC**,** 1993]. The guidelines of United Nations (2006) for the evaluation of acute toxicity, propose substantially short term (hours to few days) lethal assays, which measure LC50 or EC50. The guidelines for the assessment of chronic toxicity have a different approach, being the length of the assay related to the life cycle of the organism (generally days, weeks, months or even years). In this case, sublethal endpoints are often preferred to mortality, and NOEC used instead of LC50.

For identification of the assays that make up the individual batteries, priority is given to those for which there are methodological protocols (standards). In the Annex (section 4), a large number of bioassays for which there are standards and/or guidance documents available are listed. *In vivo* bioassay protocols have much in common with traditional toxicity test protocols, developed for chemical regulation purposes (such as the *Daphnia magna* test). In many cases the same test protocols can actually be used although, for chronic assays, the feeding and test-solution renewal schedule also may need to be adjusted. To study effects from environmental samples containing a complex mixture of substances, dilution series are frequently not used. In general the level of contamination in surface water is low. Monitoring of environmental toxicity in the range of sub-acute or chronic effects is very laborious and therefore expensive, because of the long time it can take to detect any effect.Instead, as was mentioned in 4.1.1., there is normally a need to perform sample pretreatment in order to obtain sufficient sensitivity to use short term *in vivo* bioassays, especially for surface water testing. In order to examine trends in toxicity or to detect the level of toxicity, the surface water should be extracted or concentrated. These methods will not concentrate all substances in an equal way and especially metals will fail to be concentrated. Also the bioavailability of different substances can change. This matter is further discussed in the Annex (section 3).

The monitoring and assessment of sediment quality using effect-based methods are of increasing importance. Bioassays are also mentioned in CIS guidance no. 27 as a potential tier 2 step in the evaluation of sediment. Only a few standardised *in vivo* bioassays are directly applicable to whole sediment. Instead most of the assays/toxicity tests are applied to liquid matrices obtained from the sediment, such as the elutriate and/or the pore water. A sediment elutriate is an environmental matrix that enables the replication of sediment mobilisation phenomena and the prediction of the release of contaminants from the sediment to the water column. It was first developed to evaluate the potential effects of disposing dredged material in open water and is nowadays also applied to the quality evaluation of *in situ* sediment (Arizzi Novelli et al.,2005). The 1:4 sediment:water ratio, suggested by USEPA (1991), is the most commonly employed sediment: water proportion.

Besides preparation of a sample, sampling and test conditions used for bioassay analyses may change the bioavailability of the compounds as well. The impacts of confounding factors on baseline data and bioassay responses should be well established in order to distinguish between natural variability and contamination induced stress. For a few routinely used *in vivo* bioassays for marine and freshwater organisms Postma et al. (2002) have set criteria for confounding factors based on the species ecological range.

4.4 Biomarkers

4.4.1 Biomarkers – general technical considerations

As could be concluded in chapter 2, biomarkers are already for example included in the monitoring programmes of Regional Seas Conventions, to identify the impact from substances or combinations of substances, not previously identified to be of concern, to study trends and to identify regions of decreased environmental quality.

Contrary to bioassays but similar to the ecological/community based tools (further described in chapter 5), biomarkers are analysed on field exposed, usually wild organisms. The sampling step is therefore primarily focused on the sampling of the organisms that are to be examined.

It is important to discover effects related to chemical substances before significant effects on population level can be observed. Damage at the population and ecosystem level can take a long time to repair. For certain trophic levels, recolonisation can for example take much longer time than the time frames (6 year management cycles) considered in the WFD. Ecological tools/indices (see chapter 5) are not predictive, whereas several biomarkers can detect effects caused by chemical substances at an earlier stage.

Biomarkers are frequently divided into two different categories, depending on the number of substances/groups of substances they are known to respond to:

- General (integrative) biomarkers that respond to several classes of toxic substances and, frequently also to other types of stressors
- Specific biomarkers that respond primarily to only a few /groups of/ substances.

Imposex is considered to be a very specific effect, responding primarily to organic tin compounds such as TBT, whereas lysosomal stability is a more general biomarker. Both general and specific biomarkers can be useful, depending on the monitoring purpose and prior knowledge regarding type of contaminants present at the given location. Because general biomarkers respond to several classes of compounds, they cover more substances and are therefore valuable in identifying areas of concern in environments exposed to complex exposures. Specific biomarkers are valuable in for example 2nd tier assessments to confirm field responses due to certain types of substances that are present in elevated concentrations, see e.g. case study “Monitoring imposex on water body level”.

Biomarkers can also be divided into:

- Exposure biomarkers, measuring e.g. alterations at molecular and cellular levels, including the induction or inhibition of specific enzymes involved in biotransformation and detoxification mechanisms
- Effect biomarkers, measuring responses that are highly relevant to the organism health and/or possibility to survive and reproduce, suggesting risks to higher organisational levels than individual levels

Exposure biomarkers, such as EROD[56], can provide a sensitive indication of early changes, which often represent the first warning signals of environmental disturbance. EROD can be used to detect exposure to classes of organic pollutants such as co-planar PCBs, polyaromatic hydrocarbons, planar dibenzodioxines (CCD) and dibenzofurans (CDF). Metallothioneins, peroxisomal enzymes (e.g. acyl CoA oxidase) and inhibition of acetylcholinesterase activity are other more or less specific exposure responses towards trace metals, organic chemicals and organophosphate pesticides, respectively. The advantages of biomarkers of exposure are their early response and their specificity toward specific classes of toxicants, but they do not necessarily reflect the onset of adverse health effects.

Effect biomarkers on the other hand, indicate the occurrence of various forms of molecular to cellular/tissue alterations, although the health related effects may differ in terms of toxicological and ecological relevance. Some effect biomarkers are detecting effects at early stages (such as genetic changes), whereas others are rather related to late stages (such as imposex) from a population risk perspective.

There is no strict line between these different classes of biomarkers, but a scale going from specific to general responses, from low to high ecological relevance, and from early to late responses. There are remarkably few established specific biomarkers available that primarily respond to a certain substance or group of chemical substances. During many years, several biomarkers such as EROD, metallothionein (MT), peroxisomal enzymes (i.e. acyl CoA oxidase) and activity of acetylcholinesterase (AchE) were for example considered specific to planar organic compounds, trace metals, organic chemicals and organophosphates pesticides, respectively. However, with an increase in scientific knowledge, it can be concluded that these biomarkers are not entirely specific for these compounds. For example, imidazole pesticides such as prochloraz are able to induce EROD activity according to an unknown mechanism (Sanchez et al. 2008b). Similarly, many chemicals that can induce oxidative stress are metallothionein inducers (Gagné et al. 2008). To study the effects from certain types of substances (having similar mode of actions), a combination of several different types of biomarkers could provide important information, because of the relationship between early responses of organisms and individual or population disturbances. A significant vitellogenin induction should for example trigger an analysis of intersex. Similar relationships could be observed for other relevant biomarkers such as acetylcholinesterase and parameters associated to central physiological functions (e.g. reproduction, immunity or energy).

The relationship between a biomarker response and chemical exposure is not necessarily strictly linear due to adaptive mechanisms (Mayer et al. 1992) or transient responses, as reported for antioxidant parameters (Sanchez et al. 2005). Biomarkers are studied on organisms that were exposed in the field and not controlled laboratory conditions. Even the most specific biomarkers can therefore also respond to other types of environmental stress factors, such as hypoxia or temperature increase as well as parasitic infections. However, such factors can usually be taken into account in the design of the programme and during interpretation. Another aspect to be aware of is that simultaneously acting chemicals having antagonistic effects could cancel out the response observed but still put the organism under stress. As an example, it has been proposed that coexposure to planar

[56] EROD induction indicates biotransformation activity of the cytochrome P 450-dependent monooxygenase involved in phase I of the biotransformation system

compounds that interact with the AH receptor can inhibit VTG synthesis through increased metabolism (see also fact sheets on VTG and EROD).

Broeg et al (2005) concluded that a multi-biomarker approach, based on a combination of several kinds of biomarkers is a useful prerequisite to assess the impact of environmental contamination at different levels of organisation. Hence, the application of a set of biomarkers based on complementary parameter measurements appears as a valuable way to differentiate clean and polluted sites or to describe accurately contamination effects on organisms (Flammarion et al. 2002; Galloway et al. 2004; Sanchez et al. 2008a).
See also e.g. case studies "Deployment of a multi-biomarker approach to identify the origin of wild fish abnormalities reported in a French stream receiving urban and industrial effluents » and « Swedish national monitoring programme of fish health". In some cases though, a more narrow spectrum of biomarkers are sufficient for the purpose of the study, see e.g. "Endocrine Disruptors in the Irish Aquatic Environment" and "Monitoring imposex on water body level".

Tools that respond to the cumulative exposure to several classes of contaminants are valuable to identify risks at an early stage in areas that are located far from known sources, see e.g. case study "Swedish national monitoring programme of fish health". Because the stations of such monitoring programmes are located sometimes even in pristine areas, it is vital to include tools that are very sensitive in order to be able to detect any changes at an early stage and again to cover a broad response range. Several sensitive biomarker batteries have been used on a regular basis as early warning systems for several decades.

The difficulty to link chemical exposure and biochemical response is increased by pollutant interactions during exposure to complex mixtures. In some cases regional studies upstreams using the same set of biomarkers could be sufficient to locate sources and undertake measures without actually identifying the specific substances involved. Biomarkers using sessile organisms such as mussels or *in situ* bioassays can also be included in such gradient studies. If certain combinations of biomarkers[57] are responding, this may be sufficient to identify substances that are of primary interest and support implementing suitable control measures. By studying the same mode of action using *in vitro* systems on samples taken from the environment (see Annex section 7) subsequent EDA analyses could be performed to investigate the identity of toxicants. However, if significant effects are primarily observed in a « pristine » area (located far from known local sources) it will probably be a major challenge to identify causes and suitable control measures. The reasons could be related to substances subject to a combination of long range water and air transport. Warning signals would then need to trigger investigative projects, of different character depending on the type of signals and case-specific circumstances.

Furthermore, although specific biomarkers can be used to identify early responses from certain compounds in the environment, in a WFD context, it should be kept in mind that the protection objectives when it comes to priority and other chemical substances also include other than aquatic organisms (such as predatory birds and mammals as well as humans). The effects from substances that are subject to biomagnification are likely to occur at an earlier stage in predators than in aquatic organisms such as fish. Moreover, it is also

[57] EROD induction combined with Liver Somatic Index (LSI) response would e.g. suggest that one should suspect effects from AH receptor inducers such as PAHs, planar PCBs and PCDD/Fs

possible that even if pelagic and benthic organisms could be considered equally sensitive to the same concentration of a substance, benthic organisms are generally exposed to higher concentrations of accumulating substances. If effects are studied only on pelagic fish, effects in benthic organisms[58] might be overlooked. To predict effects from substances that are subject to biomagnification, biomarkers are primarily suitable to be used in a supplementary manner to chemical analysis (including the sampled tissues), Care should also be taken to include not only different types of biomarkers in a battery but preferably also both pelagic and benthic organisms.

4.4.2 Biomarkers available

Several biomarkers are well described in scientific literature and some of them are included to assess the quality of aquatic environment in various environmental monitoring programmes. A summary of the established biomarkers for environmental monitoring purposes is given by Viarengo et al (2007).

In a more general sense, biomarkers should be reliable, relatively cheap, easy to perform, methods standardised, assessment criteria defined, and intercalibration procedures in place (see also Annex section 5). Moreover, the use of non-invasive or non-destructive methods can facilitate certain types of applications. The biomarker response should be sensitive to xenobiotic exposure and/or effects to serve as an early warning parameter. Moreover, the temporal response profiles of biomarkers after exposure to chemicals should also be known for a better understanding of biomarker results (Wu et al. 2005). The impacts of confounding factors on baseline data and biomarker responses should be well established in order to distinguish between natural variability and pollution-induced stress. For this purpose, biology and physiology of selected organisms should be known to minimise variation sources (e.g. age, gender, reproductive status). Also the mechanisms supporting the relationships between biological responses used as biomarker and pollutant exposure should be defined, as well as the relationships between biomarker responses and impact to the organisms should be clarified.

Table 4.2. lists a number of biomarkers that are used more or less frequently within regular monitoring and for which there are fact sheets in the Annex (section 6) to this report. The fact sheets provide some further practical information (such as suitable season for sampling, tissues frequently investigated, amount of sample needed and costs) but also some information about which category (exposure/effect and specific/general) the biomarker generally is considered to belong to. If other factors than chemicals can influence observed responses these are also indicated. These fact sheets are intended to provide some guidance on the usefulness and limitations related to aspects that need to be assessed before considering including them in a monitoring programme for particular purposes.

Table 4.2. Biomarkers for which there are fact sheets included in this report, with short descriptions about the mode of action studied and the types of contaminants they respond to, availability of marine assessment criteria and current status within the integrated

[58] Especially e.g. infaunal organisms, being exposed to pore water, and – depending on feeding behaviour – possibly also sediment bound material.

monitoring approach proposed by ICES as well as indicators within the Regional Seas Conventions.

Biomarker	Description	Responds to	Marine assessment criteria available (ICES)	Integrated monitoring component (ICES)	Indicator (Regional Seas Conventions)
EROD activity	Biotransformation enzyme induced by planar hydrocarbon	PCBs, PAHs and dioxin-like compounds	BAC	Core in fish	OSPAR cand
Acetylcholinesterase activity (AChE)	Enzyme implicated in nervous transmission	Organophosphates, carbamates and similar molecules	BAC and EAC (both mussels and fish)	Core in fish and mussels	
Vitellogenin (VTG) in male fish	A precursor of egg yolk, normally synthesized by female fish	Oestrogenic endocrine disrupting compounds	BAC	Core in fish	
Metallothionein (MT)	Metal scavenger implicated in protection against oxidative stress	Heavy metals and inducer of oxidative stress	BAC (mussels only)	Additional in mussels	
Amino-levulinic acid deshydratase (ALAD)	Enzyme implicated in amino-acid metabolism	Lead exposure	NO	NO	
Lysosomal stability	General health, lysosomes play a key role in liver injury caused by various xenobiotics	Several classes of pollutants, including PAH, inducer of oxidative stress, metals, organochlorines	BAC and EAC[59]	Additional in fish, core in mussels	OSPAR cand, HELCOM preCore
DNA adducts	Alteration of DNA structure able to disturb DNA function	Genotoxic compounds including PAHs and other synthetic organic	BAC and EAC	Additional in fish	
Imposex biomarkers (e.g. VDSI)	Imposition of male sex characteristics on females	TBT	BAC and EAC	Core method in gastropods	OSPAR Common, HELCOM Core
PAH bile metabolites	PAH metabolites in bile/urine represent the final stage of the biotransformation process	Indirect indicator of PAH exposure	BAC and EAC	Core in fish	OSPAR cand, HELCOM Core
Liver histopathology	General indication about liver damage, but can be diagnostic	PAHs	BAC and EAC	Core in fish	HELCOM preCore[60]

[59] Values do not differ between species but between method (neutral red retention and cytochemical respectively)
[60] Possibly as part of FDI (Fish Disease Index)

	depending on the type of lesion				
Macroscopic liver neoplasms	Visible fish liver tumors	Cancer inducing substances; PAHs	BAC and EAC	Core in fish	HELCOM preCore[61]
Externally visible fish diseases	Overall organism health External investigations of fish ; significant changes indicate chronic stress	Several classes of pollutants, incl pathogens	BAC and EAC	Core in fish	OSPAR cand, HELCOM preCore[62]
Reproductive success in (viviparous) eelpout	The females give birth to living larvae and the species is narrowly territorial ; reproductive success and embryo malformations studied ; Reproductive success is directly related to expected effects on population level	Responds to several different types of xenobiotics, including organochlorines, pesticides, PAH, metals	BAC and EAC	Additional in fish	HELCOM preCore
Intersex	Presence of ovarian tissue in male gonads compromising reproductive capacity	Oestrogenic endocrine disrupting compounds	BAC	Core in fish	
Micronucleus	Damage to genetic material of organisms ; could affect their health and potentially also their offspring.	Substances causing permanent and hereditary double DNA strand breaks	BAC	Additional in fish	OSPAR cand HELCOM preCore
Amphipod embryo alterations	Embryo malformations studied (viviparous organisms)	Overall organism health ; strong correlation observed between malformed embryos and concentrations of metals and organic compounds	BAC and EAC	NO	
Stress proteins	Early stage effects, including oxidative stress	Responds to many types of stress factors	NO	NO	
Benthic diatom malformations	Malformations; overall organism health	Significant response to metals and several pesticides, but less to other priority substances	not relevant[63]	Not relevant	
Comet assay	Sensitive tool to detect genetic damage	Substances causing DNA strand breaks	BAC	Core in fish ; additional in mussels	

[61] Possibly as part of FDI (Fish Disease Index)
[62] Possibly as part of FDI (Fish Disease Index)
[63] Limnic biomarker

Mussel histopathology (gametogenesis)	Histological studies of e.g. digestive gland and tube	Many groups of substances, including PAHs, PCBs and heavy metals	BAC and EAC	Core in mussels	
Stress on stress	Survival in air	Many groups of substances, including crude oil, copper ions and PCB	BAC and EAC	Core in mussels	
Scope for Growth	Measures alterations in the energy available for growth and reproduction.	Many groups of substances, including DEHP, aromatics, PCP, copper, TBT and dichlorvos	BAC and EAC	Additional in mussels	

However, there are many additional biomarkers available, and new being developed. A research area of interest is the development of molecular biomarkers using different types of OMICS approaches, see chapter 7 and examples provided in the Annex (section 10) to this report.

4.4.3 Biomarkers to study endocrine disruption – WFD aspects

Endocrine disruption due to xenobiotic substances is a major concern, because of the severe effects that can occur on population levels. Endocrine disruption could be related to substances having several different types of modes of action. Well known historic cases of population level effects are related to DDT and PCB, substances subject to biomagnification and considered to have been involved in reduced bird and seal populations respectively. Also the effects from organic tin compounds, not being subject to biomagnification have caused extinction of several gastropod populations. DDT, TBT and planar PCBs are included in Directive 2008/105/EC and should be considered in chemical status classification.

The chemical analysis of TBT in water is difficult because of usually highly variable concentrations and high detection limits. In marine sediment, the substance is still frequently found at very high concentrations compared to recalculated values based on water-EQS, but there are several uncertainties involved both in this recalculation and in assessing bioavailability. Imposex biomarkers in gastropods are frequently used to investigate effects related to tributyl tin compounds and required within OSPAR. If other possible factors (such as parasites) can be ruled out, it is reasonable to assume that if surface sediment concentrations are elevated[64] and significant levels of imposex are observed, the TBT is likely to be of concern (see also the case study “Monitoring imposex on water body level”). However, even if not observing imposex, the TBT concentrations could be above the EQS and thus pose a threat to other, more sensitive species[65].

[64] Imposex is an irreversible effect, and although maybe of less importance to consider in an assessment of ecological status, this could be a critical aspect in the assessment of chemical status, if analysing the effects on organisms having a long life span.

[65] Not all gastropod populations have re-emerged after the major extinctions in the late 60s, early 70s. It is also possible that certain populations may develop resistance.

As previously described, much concern is now also given to the presence of oestrogenic substances, including alkyl phenols and certain highly oestrogenic pharmaceuticals, in surface waters. Oestrogenic substances of different types and origins are frequently found in effluents from sewage treatment plants (STP) and also other sources. The treated effluents from these STPs are directly discharged into surface waters and the corresponding water body could be at risk. To investigate further, the biomarker vitellogenin (VTG) in fish could be monitored and examined, preferably along with results from determination of intersex and sex ratios. VTG and intersex are very sensitive effect-based tools that respond to exposure to oestrogenic substances. VTG is considered an exposure biomarker, whereas intersex an effect[66] biomarker but both can be considered "early warning" tools. If for example, no or very low VTG induction and no significant prevalence of intersex is detected, compared to baselines, further monitoring of oestrogenic substances would be less prioritised. To illustrate, see case study "Estimation of Oestrogenic Compounds in Irish Surface and Waste Waters" (Annex section 1) and for more information about the biomarkers VTG and intersex, fact sheets in Annex section 6.

4.4.4 Biomarkers specific to the limnic environment

In the WFD context, it is of particular interest to identify biomarkers that could be used also in limnic environments. Although the biomarkers described so far have previously been used on a regular basis primarily in the marine environment, many can be applied also in the limnic environment (Sternbeck et al 2008). However, the baseline can be expected to be different for several of these tools, due to the use of other species for example. The assessment criteria developed for marine use may therefore not be valid in the limnic environment.

There are also biomarkers that are only used in the limnic environment, usually also relevant primarily to either lentic or lotic habitats.

A limnic biomarker tool that is gaining in popularity in some Member States is the monitoring of malformed diatoms. One contributing factor is that the additional costs to also include the assessment of malformations are low if samples are anyway collected and analysed to investigate effects from eutrophication. Besides studying the endpoint malformed individuals (frequency), it is suggested that other community level variables should be studied in order to facilitate data interpretation. Thus, studying such impact on diatoms could be considered both a biomarker and community based tool ("ecological indicator", see chapter 5). The tool is shortly described in the fact sheet in the Annex (section 6) and references therein.

Another limnic biomarker that would be possible to asses in a coordinated programme of both eutrophication and effects from hazardous substances is mentum deformations in chironomids. This analysis has been in use for several decades in Sweden and the Netherlands but some validation studies would be necessary.

[66] Intersex can have a later deleterious impact on reproduction

4.5 Integrated assessment of results

Total monitoring costs can be reduced if sampling is coordinated with chemical analytical programmes. If biota is monitored, adding biomarker analyses do not necessarily add substantially to the total costs, although there are also practical issues to be aware of[67]. Integrated monitoring approaches also facilitate integrated or stepwise interpretations of the results, facilitating the assessment of environmental quality and providing a better decision support in managing risks. *In vitro* assays can usually be applied to any type of sample, but one needs to consider whether it is likely to find chemicals possessing the mode of action studied in the compartment sampled, and whether they are expected to respond to the substances of concern[68].

4.5.1 Evaluating effect-based data

Difficulty in analysing biomarker and bioassay responses by environmental managers has been identified as a major obstacle to large-scale deployment of effect-based monitoring tools. A common approach is to evaluate such data in a relative way – either in time (trend analysis) or between investigated areas, but now also criteria expressed as absolute values are available. For many of the effect-based monitoring tools, an integrated evaluation approach is often considered the most appropriate.

4.5.1.1 Criteria for individual endpoints, to evaluate ”status”

ICES has developed assessment criteria for several effect-based tools (biomarkers and bioassays), see table in Annex (section 8). Most of the criteria are based on relative comparisons between baseline levels “background response” and levels observed at impacted sites (BAC, Background Assessment Criteria). However, there are also EAC (Environmental Assessment Criteria) values, representing levels of response below which unacceptable responses at higher, such as organism or population, levels would not be expected. However, these are applicable only to *in vivo* bioassays and effect biomarkers but not exposure biomarkers. For detailed information on how the BAC and EAC values were developed for individual biomarkers and bioassays, recent reports by ICES should be consulted (e.g. Davies & Veethak 2012).

4.5.1.2 Integrated or stepwise approaches – to assess quality and identify the cause

Due to the large number of pollutants encountered in the aquatic environment and the various effects of these pollutants, as was pointed out above, biomarkers and other effect-based tools are preferably used in a battery (see e.g. Annex section 2). The data should preferably also be evaluated in an integrated and/or stepwise manner, especially if the purpose is to assess the overall risks on population and community levels, thus assessing the overall status, rather than primarily assessing trends of individual variables. A Dutch suggestion on how to interpret *in vivo* bioassay results has recently been published (Struijs et al 2010; see also case study “Monitoring concentrated surface water with in vivo bioassays in the Netherlands”).

[67] Such as the availability of sufficient material and time of the year to sample; see fact sheets in the Annex.

[68] The application of *in vitro* assays to detect oestrogenic substances is expected to be relevant primarily for water samples (and perhaps sediment) but could be less suitable for biota samples (because of the potential impact from endogenic hormones and not primarily xenoestrogens)

Stepwise approaches to the evaluation of effect-based data also facilitate identifying suitable management options. If for example, *in vitro* assay data suggest a certain effect to occur, a confirmation using biomarkers investigating the same mode of action may be necessary. If effects are indicated by biomarker observations in the field, it may be necessary to identify the causing substances in order to implement suitable measures (such as regulating the use of certain substances). As an example, integrated monitoring of Scope for Growth and chemical contaminants in mussels has been used successfully to detect, quantify and identify potential causes of pollution (reviewed by Widdows & Donkin 1992). Specific biomarkers suggest already the type of compounds involved and to study further. If less specific but at least indicating a certain mode of action, one approach could be to try to confirm the presence of substances having the same mode of action in samples taken from the investigated area such as gradients from potential point sources, by using *in vitro* assays (see e.g. Annex, section 7). Also WEA performed retroactively could provide valuable information in such a context.

Hamers et al (2010, 2013) suggest that "Toxicity profiles" (also called "toxicological fingerprints"), could be useful in environmental quality assessments in the following ways:

- translation into hazard profiles. The relative distance to an adequate baseline toxicity profile (such as toxicity profiles from reference locations), reflect the desired or acceptable environmental quality
- translation into risk profiles, based on the ratio between the actual bioassay response and a bioassay response considered to be safe for environmental health
- selecting samples with relatively high toxic potency for further identification of causative compounds using in depth effect-directed analysis (EDA) strategies (chapter 7).

Toxicity profiling require that the effects can be studied using a high throughput approach (*in vitro* or *in vivo* bioassays focusing on different modes of action after preconcentration of samples).

4.5.1.3 Integrated assessment of biomarker data

Several authors have developed integrative indices able to summarise responses of a set of biomarkers into a single value and/or a graph (Narbonne et al. 1999; Beliaeff et Burgeot 2002; Chèvre et al. 2003; Broeg et al. 2005). Among these indexes, the "Integrated Biological Responses" (IBR) described by Beliaeff and Burgeot (2002) is frequently used in field and laboratory studies to provide an integrated view of multi-biomarker responses (Arzate-Cardenas & Martinez-Jeronimo 2011; Serafim et al. 2011). This index calculates a value providing a global response characteristic of the impact of environmental stress on the basis of multi-biomarker responses, and a star plot (fig 4.2.) representing individual responses of selected biomarkers. Several modifications of this tool were performed to correct any inconvenience such as the lack of consideration of biomarkers that can be induced and inhibited or the variance of index value (star plot area) with respect to biomarker position (Sanchez 2007).

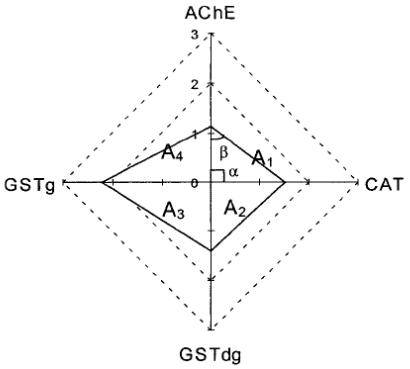


Figure 4.2. Example of a biomarker star plot calculated for a single station. AChE = acetylcholinesterase, CAT = catalase, GST = glutathione-S-transferase in mussel digestive gland (dg) or gills (g).

More recently, (Dagnino et al. 2007) proposed a decision-support system also named «expert system» that utilises a large set of biomarkers measured in marine mussels to translate complex biological responses into a relatively simple, easy to understand and objective evaluation of the changes in the organism physiology induced by pollutants. This tool proposed a classification scale that considers the various characteristics of the biological responses to environmental stressors at different levels of biological organisation. It was developed and calibrated utilising large data sets of biomarker responses measured in field and mesocosm studies and accumulated over the past two decades. A score represents the health status of organisms and allows to predict adverse effects. The software of the expert system appears as a tool easy to use by scientists and environmental managers. However, the transfer to other sentinel species and the additional integration of novel biomarker requires a large data set.

Fish biomarkers can be divided into a few major categories, such as reproduction, condition and metabolism, liver function and immune response. It is reasonable to assume that biomarkers related to reproduction are of the highest relevance from an ecological perspective. However, among the biomarkers that can be related to disturbances to the reproductive system, certain biomarkers are of higher relevance. A Swedish draft proposal on a weight of evidence approach to be used for an integrated assessment of observed responses from fish biomarker batteries generated by the Swedish marine fish monitoring programme is included in one of the case studies in the Annex (section 1) of this report (“Swedish national monitoring programme of fish health”) also presented. The level of response observed in each biomarker is first assessed individually. The weight of each biomarker is then related to both the degree of ecological relevance and the type of function it is related to, before the final index will be calculated, indicating overall health status of the fish. The assessment scheme proposed is to be used for biomarker batteries, but does not necessarily imply that all biomarkers need to be included in each assessment, but the number and choice of studied biomarkers are allowed to differ between locations.

4.5.2 Integrated assessment of chemicals and effects

A key advantage of measuring chemical concentrations in environmental media such as water, sediment and/or biota is that these provide evidence of past and/or present exposure to the analysed contaminants. Moreover, when long time series of data exist, the trends can be used to assess the decrease or increase of risks and the necessity to regulate emissions of certain individual substances. Chemical data from several trophic levels also provide information on persistence, bioavailability and accumulation in food chains and thus risks related to human consumption and top predators.

However, chemical characterisation by itself does not provide specific biological information about potential hazards to organisms. By truly integrated monitoring strategies, involving chemical analyses, ecotoxicological tools and the study of population/community responses in the same water body, and preferably time of the year and even on the same populations and individuals, a holistic picture can be obtained. Such an approach can tell whether the chemicals analysed are bioavailable and giving rise to negative health effects in aquatic organisms, and whether population effects are observed.

The concept of weight of evidence (WOE) refers to the integration of data generated within such a multidisciplinary approach including data from different studies, or lines of evidence (LOEs). These LOEs address questions relating to the presence and biological effects of chemical pollutants: traditional chemical analyses are combined with laboratory and field-based studies to assess the bioavailability of pollutants to selected species, and the onset of adverse effects at different levels of biological organisation, i.e. from molecular, to organism up to community level.

WOE methods are often key components of Ecological Risk Assessment (ERA) and also in line with the European Water Framework Directive 2000/60/CE which requires member states to evaluate and classify the ecological status of water bodies integrating different quality elements, although a one out - all out approach is used to assess chemical data[69]. An integrated monitoring approach and integrated evaluation of data, based on both chemical and effect-based tools, would not only provide more holistic status classifications but also better decision support within water management. This would greatly benefit prioritisation issues, especially if measures are costly.

The combination of multiple LOEs represents an added value to monitoring and management protocols, especially compared to regulatory frameworks which still rely on chemical characterisation relative to Quality Guidelines (QG) as stand-alone decision criteria.

The WOE approaches can use either qualitative or quantitative assessments to set individual LOEs in an integrated assessment of impairment or risk. The simplest methods are qualitative interpretations of different results, while quantitative approaches are based on more structured mathematical and statistical elaborations, providing indices for each LOE and for their overall integration. Although there is not a standardised procedure, the comparison and aggregation of heterogeneous data by quantitative methods can rely on the assignment of weights, thresholds, indexing criteria, classification of endpoints, comparison to reference conditions, normalisation functions, and identification of impairment classes.

[69] If concentrations of an individual substance exceed the corresponding EQS of that substance, chemical status is not good. If the EQS of a RBSP is exceeded, the ecological status is considered “moderate”.

The standardisation of similar procedures requires initial assumptions and expert judgments which should be supported by a robust scientific rationale. WOE approaches are further implemented by the multi criteria decision analyses which, beside scientific evidence on sediment quality, often consider also other relevant, i.e. social or economic criteria. These methods reveal similarities or potential conflicts among stakeholders and experts with different views, allowing decision-makers to prioritise best management options.

WOE models have been recently presented to integrate various lines of evidence like chemical, physicochemical, ecotoxicological and biological data [Dagnino et al., 2008; Semenzin et al., 2008; Benedetti et al., 2011; Micheletti et al., 2011; Piva et al., 2011]. Logical flow charts and mathematical models were also defined for each LOE with some assumptions based on expert evaluations to elaboration results in synthetic indices specific for various LOEs and a final classification of overall hazard or quality [Benedetti et al., 2011; Piva et al., 2011].

In several studies comprehensive investigations using Weight of evidence approaches and triad approaches have been applied (for review eg, Chapman & Hollert 2006). Within contaminated sediment management (remediation and dredging) the TRIAD approach has been in use for several decades. In Germany, recommendations were made for the use of an integrated stepwise approach combining toxicological, chemical and ecological information to assess and evaluate the quality of sediments (Ahlf et al. 2002a, b). In the marine context, integrated monitoring and assessment approaches have been suggested by both HELCOM (the CHASE tool) and OSPAR (traffic light system[70]) and a recent publication by ICES describes an integrated assessment framework for contaminants and effects (Davies & Veethak 2012). The strategy on integrated ecosystem assessment is based on sediment monitoring (chemistry, characteristics, bioassays, benthic ecology), water monitoring (passive samplers, bioassays, hydrography, bioassays, water chemistryetc) and biota monitoring (tissue chemistry, fish biological effects, mussel biological effects, gastropod biological effects). A NORMAN Protocol for integration of biological and chemical test methods is available on the NORMAN website[71].

4.6 Toward a suitable WFD and MSFD toolbox

The main purposes to add effect-based tools in aquatic monitoring programmes, are not to replace chemical analysis but rather to take into account the effects of substances that are not normally or easily monitored chemically (“non listed substances”, substances with high detection/quantification limits, intermittent exposures/variable concentrations, emerging contaminants). Furthermore the use of effect-based tools helps to obtain an approach that considers cumulative/mixture effects.

The most suitable choice of established effect-based monitoring tools (biomarkers and bioassays) to use, in particular within the WFD, but also MSFD context, will depend on the purpose and situation. Chemical monitoring is a key part of the WFD monitoring programmes to assess water quality. Chemical analysis is indeed necessary in identifying

[70] By first evaluating each individual substance and effect observed against the established BAC and/or EAC value.
[71] http://www.norman-network.net/sites/default/files/files/QA-QC%20Issues/protocol_v1_1b_version_15oct.pdf

for example persistent and bioaccumulable substances in biota to identify also possible risks for human health, top predators and aquatic organisms in high trophic levels.

If water is monitored, bioassay batteries performed on the same samples (or at least taken from the same location and at the same time) could offer practical, complementary tools to cover additional substances with a certain mode of action but that are not analysed chemically. Table 4.1. includes a few of the established *in vitro* assays, considered suitable to investigate surface waters and responding to WFD relevant mode of actions, such as endocrine disrupting compounds (affecting oestrogen, androgen and thyroid receptors), substances binding to the Ah (« dioxin ») receptor and genotoxic compounds (cf list of « main pollutants » in Annex VIII to WFD, point 4). Additional *in vitro* tools responding to Ah receptor activation, sex hormone disruption, thyroid hormone disruption, genotoxicity, oxidative stress, neurotoxicity, immunotoxicity and cytotoxicity, are listed in the Annex section 7. Chapter 4.2.3. and Annex section 2 also include proposed approaches to investigate oestrogenic potential of environmental samples.

After preconcentration of water samples, short term *in vivo* bioassays (using WFD relevant endpoints) have also proven valuable in this context to identify areas of elevated risks. Because *in vivo* bioassays are generally developed from corresponding toxicity tests (used within chemicals testing), a large number of standards are available (cf Annex section 4). A novel water sampling approach is based on passive sampling, a tool that is applicable not only to take samples for chemical analyses but also bioassays (Annex section 3).

The identification of individual, local sources, by backtracking (sampling in a gradient from suspected sources and the emissions from these) is possible also when using bioassays. To identify the major substances that are likely to cause the observed effects in a sample are also possible by using EDA/TIE approaches (further described in Chapter 6).

Marine monitoring programmes are frequently focused on sediment and biota sampling, to quantify levels of persistent and/or bioaccumulable substances (cf established monitoring programmes within the regional sea conventions), but there are also biota monitoring programmes in lakes. The WFD requires trend monitoring in sediment and/or biota of accumulating substances. With the introduction of additional biota standards (cf chapter 2.3.), there will be an increase of efforts for biota monitoring. If biota is anyway sampled for chemical analysis, adding a selected number of biomarker analyses can provide a cost effective, integrated approach. There are a large number of established biomarkers that cover both WFD relevant mode of actions and general health effects (see e.g. table 4.2. and fact sheets in the Annex section 6). Most experience in biomarker monitoring has so far been gained in the marine environment, but most fish biomarkers can be used also in the limnic environment, although the marine assessment criteria may not be applicable (cf biomarker fact sheets, specifying limnic/marine applicability; and marine assessment criteria in Annex section 6 and 8 respectively). There are also some new methods that show promise for future use in the limnic environment, such as diatom malformation. Again, there are also numerous in vitro tools, covering WFD relevant mode of actions and that could be used to analyse biota. Ah receptor binding *in vitro* bioassays are already accepted as screening tools within food legislation (criteria for dioxin and dioxin-like compounds), and the major reason to use such a tool in that context is the significantly lower cost.

Although less straightforward than in a bioassay approach, biota/biomarker sampling in a gradient can be performed to identify individual, local sources, using either sessile

organisms or cage experiments. If specific biomarkers are included they can also provide an indication about what type of substances (and therefore sources) to suspect. However, if there are no such indications, and it is necessary to identify the main suspects, a stepwise approach can be used. In most cases it is possible to study the same mode of actions using either in vitro assays or biomarker analyses (cf Annex section 7). Once effects are confirmed to occur in corresponding in vitro (or in vivo) bioassays performed on samples from the investigated area, an EDA/TIE approach can be used.

Furthermore sediments are frequently monitored, in particular to assess trends of persistent substances but also to predict effects that could be of concern to benthic organisms. CIS 27 suggests a two tiered approach (in which effect based tools can be used in the 2nd tier) to assess sediment quality and the need for undertaking measures (such as remediation). Most in vitro bioassays are applicable also to sediments and Annex section 4 includes also in vivo bioassays developed to test whole sediment samples. It is also possible to use water based in vivo bioassays on pore water or sediment elutriates.

Annex section 8 includes a large number of marine assessment criteria, developed by ICES to be used in integrated monitoring and assessment programmes. Furthermore, in vitro assay results are frequently expressed on a chemical equivalent basis, and if chemical quality standards are available for those reference substances, it is possible to obtain an indication about the level of effects.

The MSFD requires effect based tools to be included in the assessment of status, but which tools should be used have not been specified. The work to identify such tools are ongoing within the regional sea conventions and some individual member states. The WFD does not require legally binding effect-based tools to be included in surveillance and operational monitoring programmes, but these tools could be used for investigative monitoring and could offer a valuable support in the assessment of water bodies quality.

5 ECOLOGICAL INDICATORS OF CHEMICAL POLLUTION

Ecotoxicology provides several tools to detect effects of single or mixed toxicants (on samples tested in the laboratory and on organisms exposed in the field) at individual or suborganism levels (chapter 4 on biomarkers and bioassays). The *risks* of population and community level effects can be estimated by the combination of chemical and effect based tools on lower levels of organisation. However, in the context of the WFD, population and community are the biological organisation levels on which effects should be assessed within ecological status classification (see biological quality elements specified for different types of water bodies in Annex V of the WFD). Therefore, individual biomarkers can, by definition, not be used as biological quality elements, except if endpoints on higher biological organisation levels are also added (see e.g. malformed diatoms for which community level variables are suggested in the evaluation, shortly described in chapter 4 and fact sheet in the Annex, section 6).
According to article 2 of the WFD, the definition of "Ecological status" is: "an expression of the quality of the structure and functioning of aquatic ecosystems associated with surface waters, classified in accordance with Annex V". The biological quality elements of rivers, lakes, transitional waters and coastal waters defined in Annex V of the WFD, all include composition and abundance of aquatic flora and composition and abundance of invertebrate fauna. For lakes, coastal and transitional waters, also biomass of phytoplankton is included and rivers, lakes and transitional waters also include composition and abundance of fish fauna. For rivers and lakes, age structure of fish fauna is also included. With these biological quality elements, the classification will largely be based on community structure rather than function, although the definition of ecological status in the WFD clearly also includes function.

The assessment criteria (primarily based on values of different indices) for the biological quality indicators do not respond well to the effects from hazardous substances[72]. Specific tools for the assessment of hazardous substances applied in the context of WFD monitoring are extremely rare (Birk et al, 2012). Therefore, ecological status related to hazardous substances, is at the moment more or less only based on concentrations of specific pollutants, i.e. an entirely chemical, risk based, approach. Out of 300 methods reported to be used on a European scale to assess ecological status, only four have been reported to respond to either "heavy metals" or "organic compounds" (Birk pers comm): the Infaunal Quality Index (IQI; UK), the Danish Quality Index Ver2, the Spanish Multivariate AZTI Marine Biotic Index (M-AMBI (Spain; Muxika et al. 2007) and the French "Benthic Opportunistic Annelida Amphipoda Index/Benthic Opportunistic Polychaete Amphipoda Index" (Dauvin and Ruellet, 2007; 2009).

[72] For example, the sensitivity values of observed species at a particular site has a heavy impact on the value of the marine BQI (Benthic Quality Index). The sensitivity values for the same species vary between geographical areas and ranked according to presence in different types of environments. Species that are common in areas with low number of species obtain a low sensitivity value. However, different species may be sensitive to one type of stress and not another type of stress. In general, the main reason for finding certain species at particular coastal sites is determined by parameters related to organic load (nutrients, oxygen levels etc) and not toxic substances. The BQI can also be applied to benthic communities of lakes and is then based on the sensitivity of different chironomid species to low oxygen levels

Nevertheless, there are also a few novel tools that could be used to directly monitor effects of stressors on community levels. They are related to impact on structure (species composition and abundance) and/or function (such as species traits). This section describes four such tools. Three of them are more specifically related to effects from hazardous substances (SPEAR index, NemaSPEAR index and PICT), while a third approach responds to and can discriminate between multiple stressors on ecosystems (a multimetric index approach based on bioecological traits of species). The ecological indicators PICT and SPEAR do not only provide information about effects occurring on community levels, but also indications about what substances or groups of substances[73] that are likely to be causing the effects. However, the ecological indicators identified in this report are only available for a limited number of applications and do not for example include effects on fish.

5.1 A multimetric index approach based on bio-ecological traits of species

A new approach within biomonitoring is to consider the ecological role of communities, based on their functional, rather than structural composition, through the identification of species traits (see e.g. Usseglio-Polatera et al., 2000). The resistance and resilience characters adopted by taxa determine the response of communities to disturbance events. Undisturbed communities display a diversity of species traits, whereas the communities downstream of a pollution source consist of those species that have the suite of traits that are tolerant to the new conditions. Those species that do not have these traits cannot survive.

The advantage of using functional traits instead of taxonomic composition (entirely structural approach) of communities is bound to the *a priori* predictable response of traits to individual stressors (Van den Brink et al., 2010; Bonada et al., 2006; Statzner et al., 2010), because each pressure affects different traits. Moreover, the response can be predicted following a mechanistic approach, i.e. considering the functional role of the trait in the organism and in the ecosystem.

A new multimetric index, based on benthic macroinvertebrate communities was developed by Archaimbault et al (2010), who described the communities in terms of 22 biological and ecological traits considered as sensitive to sediment toxicity.

Each species is in fact characterised by biological (e.g. life cycle, respiration mode, reproduction, body size, etc.) and ecological traits (e.g. feeding habits, habitat preference, tolerance to stressors, etc.) selected by evolution as strategies to cope with environmental stress (habitat templet theory) (Townsed et al., 1994). Each trait is described in multiple categories. The affinity of each macroinvertebrate taxa for the different categories of a trait is calculated using a fuzzy coding procedure (Chevene et al 1994). The fuzzy coding procedure is based on correspondence analysis and uses positive scores to describe the affinity of a species for different modalities (i.e. categories) of a given variable.

Community composition at each site is thus described as relative abundance of trait categories. A non-parametric multiple comparison statistical procedure is used to compare

[73] A significant response in SPEARpesticides would indicate that pesticides are responsible, a significant PICT signal after e.g. irgarol exposure suggests that irgarol is present at elevated concentrations, although it is also possible that co-tolerance due to other substances could be the reason. See chapter 5.

relative abundances of trait categories between groups of sites assigned to different quality classes (e.g. high versus good; good versus moderate, etc.), to identify the combinations of trait categories that best separate sites between adjacent toxic quality classes. Based on such sets of trait categories, a statistical procedure has been proposed to allocate sites to toxic quality classes from the attributes of its benthic macroinvertebrate community.

Predictions from this trait-based functional tool in French running waters achieved approximately 73% of correct site post-assignments to toxic quality classes pre-assigned using chemical criteria (i.e. the French water quality assessment system SEQ-Eau based on the presence of metals, PAH and PCB in sediment). Results indicate that the trait approach may offer an *in situ* functional tool for stream sediment contamination assessment at community level, but effort is needed to fill the gaps in species trait information.

5.2 The SPEAR index

The SPEAR (SPEcies At Risk) bioindicator system, based on biological traits, was shown to be highly sensitive to particular groups of toxicants and relatively independent of confounding factors (Beketov and Liess, 2008 and Liess et al., 2008). The index is called SPEAR index (SPEcies At Risk) and measures the proportion between sensitive (SPEAR) and less sensitive (SPEnotAR, “SPEcies not At Risk”) species, and is expressed as a percentage.

SPEAR index (%) =[{number of SPEAR}/{number of SPEnotAR}]*100

Thus, the higher the SPEAR index value, the less impacted the area is anticipated to be. The sensitivity aspects includes an assessment of both physiological sensitivity and the recovery potential, see table below.
The SPEAR index is relatively independent of abiotic environmental factors other than the selected class of organic contaminants (Liess and von der Ohe, 2005) and is applicable across different biogeographical regions (Schäfer et al., 2012).

Currently the SPEAR system includes two types of indicators designed for two different types of contaminants (particularly for two different organic toxicants):

(i) $SPEAR_{pesticides}$ designed for agricultural pesticides occurring in water in short-term pulses. To include the $SPEAR_{pesticides}$ index into monitoring programmes according to the EU Water Framework Directive the boundaries of ecological status classes for this index have been defined in small European streams (von der Ohe et al., 2007; Beketov et al., 2009). However, in large-scale river systems (e.g., in medium-size and large rivers) sensitivity, independence of confounding factors and validity of the ecological status classes’ boundaries of this index remain to be checked.
Species that are considered potentially sensitive to pesticide exposure based on an assessment of recovery potential need to fulfil criteria for all three traits, reported in the following table, in order to be further assessed.

Potentially sensitive to pesticide exposure (further assessment is done)	Classified as SPEcies NOT At Risk
Generation time exceeds 0.5 year	Generation time less than 0.5 year
Poor migration potential	Good migration potential
Aquatic larval stages during high exposures	Adult stages emerged before May (during high exposures): no aquatic exposure

In the study by Liess & von der Ohe (2005), species fulfilling the above criteria to be classified as potentially sensitive to pesticide exposure were further assessed regarding their relative physiological sensitivity. This assessment was based on the relative sensitivity observed when comparing EC50 values of the particular species to a certain substance, to the corresponding EC50 for *Daphnia magna* for the same substance. The relative sensitivity, S, is calculated by the following equation (von der Ohe & Liess 2004) [74]:

$S = \log (LC50_{Daphnia\ magna} / LC50_i)$

The obtained median relative sensitivity observed (-0.36) was used as cut off to finally identify species that should be considered SPEcies At Risk and thus included in the final calculation of the index. Thus the species are grouped according to their sensitivity to toxicants (based on the relative species sensitivity distribution rank) and their life cycle traits.

In the study by Liess and van der Ohe (2005), twenty central European streams were investigated and showed that a measured pesticide concentration of 0.01*EC50 led to a short and long term reduction of abundance and number of SPEAR and a corresponding increase in species not at risk (SPEnotAR). Even concentrations of 0.001*EC50 correlated with long term change in community composition. The SPEAR increased when there were undisturbed stream sections available upstreams, thus highlighting the need to also take conditions upstreams and recolonisation aspects into account.

(ii) $SPEAR_{organic}$ specific for organic toxicants with a relatively constant exposure regime (e.g., synthetic surfactants, petrochemicals) (Beketov and Liess, 2008). $SPEAR_{organic}$ was found to be highly dependent on organic toxicants such as synthetic surfactants and petrochemicals, and relatively independent on natural environmental factors along a large-scale river continuum. Nevertheless, further studies focused on relations of this index with both the target stressor (i.e. organic toxicants) and natural environmental factors are necessary to prove its applicability in different river systems and regions. Also validation of this index is currently not as advanced as it is for $SPEAR_{pesticides}$.
Future application of $SPEAR_{organic}$ in bioassessment should take into account possible uncertainties associated with computation of this index, which include: (i) extrapolations in deriving of taxon-specific $S_{organic}$ sensitivities (discussed in Von der

[74] The sensitivity of the species to toxic stress is in this study ranked relative the sensitivity of *Daphnia magna* for the same compound (metals and organics respectively).

Ohe and Liess, 2004), and (ii) effects of possible confounding factors on $SPEAR_{organic}$ values. The expected confounding factors are: (a) non-continuous exposure profiles with post-contamination recovery periods (the pesticide-specific SPEAR system is suggested, (Liess and von der Ohe, 2005), (b) effects of landscape factors facilitating recolonisation (e.g. upstream undisturbed area, Liess et al 2008), (c) effects of toxicants with specific receptor-mediated modes of action (e.g. neonicotinoids, Beketov and Liess, 2008b), and (d) specific sublethal effects of contaminants (e.g. drift-initiating action of neurotoxic insecticides, Beketov and Liess, 2008c). These factors can result in underestimation of the effects of organic toxicants, or hamper comparison of contaminated sites. However, all these factors are expected to influence not only $SPEAR_{organic}$, but also any of the currently applied bioassessment indices.

Using the SPEAR approach on invertebrate data from Swedish streams monitored within a national programme, the SPEAR index varied between 60-80%[75]. Chemical data are not available and the obtained data can therefore only be compared to other parameters such as percentage farmland cover. After recalculation using the new version of the index, the correlation between obtained SPEAR values and the percentage farmland cover is improved. This correlation was also observed if using PCA (Principal Component Analysis). The Swedish dataset used is from the year 2000 and the type of crops used at that time is now being investigated in order to predict retrospectively what types of pesticides were present. From a Swedish perspective, the SPEAR metric is therefore considered promising but needs to be adapted to northerly conditions regarding landscape and climate and compared to other metrics that quantify ecological change.

The SPEAR concept is applicable to assess the effects on invertebrate communities in rivers but not lakes or coastal areas and also not to temporary streams. Sampling should be performed in early summer (June, July) not too long after the main period of pesticide application. Sensitivity data and information on other relevant traits for the taxa are included within the database used for the SPEAR online calculator[76]. Validation studies were so far performed in Finland, Germany, Sweden, France, Spain, Czech Republic and Australia (Schäfer et al 2012, Wolfram et al 2012). Nonetheless there is a need for further validation before the SPEAR indices can be used on a regular basis and as part of the WFD classification. In particular, the baseline sensibility and variability of the method need to be assessed (Wolfram et al 2012). Moreover, the robusteness of the approach in being contaminant-specific needs to be proved. -

The sensitivity rank of different species is relative and specific for a certain dataset. Therefore, a species that is considered to be “SPEAR” in one particular dataset, can be considered “SPEnotAR” in another dataset depending on the relative frequency distribution. Nevertheless, the extensive validation indicates that a SPEAR index below 33% can be suggested as a sufficiently significant response to state that the site is disturbed (Beketov et al 2009).

The Spear index for pesticides has been widely applied, in different rivers, different toxic compounds and different geographical conditions, but mainly in research or risk assessent studies, not in routine monitoring. So it can be defined a «promising tool», (see also Beketov et al 2009).

[75] Willem Goedkoop pers comm
[76] http://www.systemecology.eu/SPEAR/index.php

5.2.1 The NemaSPEAR index

The Nematode Species At Risk (NemaSPEAR) index has been developed for assessment of sediment pollution, particularly in soft sediments (Höss et al 2011). Free-living nematodes occur in great diversity and at high densities in every type of sediment and they are frequently the dominant taxon in (soft) sediments that occupies key positions in benthic food webs because nematodes comprise various feeding types. Moreover, they exhibit a wide a range of sensitivities to pollutants. Accordingly, multivariate methods were used to classify nematodes as species at risk (NemaSPEAR) or not at risk (NemaSPEnotAR). The classification was derived from a large empirical dataset of field samples characterised by varying levels of contamination. Species classified as being at risk occurred only in low-level polluted samples; species not at risk were especially present in high-level polluted samples or in all samples. In addition, it was distinguished between metals and organic toxicants leading to two indices, the NemaSPEAR[%]$_{metal}$ and the NemaSPEAR[%]$_{organic}$. Similar to the SPEAR, the indices are calculated as the percentage of the abundance of nematode species at risk with respect to the abundance of all species. The indices have shown good correlations with the toxic potential of independent field samples and responded dose-dependently to chemical concentrations in two model ecosystems (Höss et al 2011). Moreover, during the derivation of the indices the influence of sediment texture could be clearly distinguished from that of the toxic potential of the sediment samples. Thus, the NemaSPEAR index is possibly a promising tool for effect-based monitoring in terms of sediment quality assessment. Nonetheless Wolfram et al (2012) found a low power of the method in discriminating between sites with different pollution levels.

5.3 Pollution-induced community tolerance (PICT)

PICT (Pollution Induced Community Tolerance) has been suggested as a sensitive tool to track changes in the community function (and therefore indicative of structural changes) that can be attributed to toxic substances. The PICT approach was developed by Blanck & Wängberg and Blanck et al (1988). It is not an index but rather a tool to explain why the community composition had to change, when exposed.

The approach relies on the assumption that sensitive components of the exposed community (species, genotypes or phenotypes) will be replaced by more tolerant ones during exposure, thus leading to an increase of community tolerance. PICT is measured with a functional test that detects the consequences of selection pressures. Tolerance development can be, for example, measured as a shift in the Effect Concentration (usually EC_{50}) that is obtained with a short-term toxicity test based on an ecophysiological endpoint. Such an endpoint is preferably related to community metabolism (photosynthesis, respiration, protein synthesis, nucleic acid synthesis etc).

The PICT-method, which uses the theoretical basis of toxicology (the dose–response model) to quantify community effects, was proposed as a tool with strong predictive ability for causative links between toxicants and their adverse ecological effects (tolerance levels). The advantage of the PICT method is that tolerance is less sensitive to natural variations of sampling sites than other community characteristics, such as microbial biomass, because it is an integrating characteristic of a community. Thus PICT provides the opportunity to isolate effects of individual stressors in systems impacted by multiple stressors, multiple contaminants as well as other anthropogenic stressors.

Because tolerance is measured, it is important that the effects observed only reflect the changes developed during the selection phase. The quantification of “average tolerance” is made by using short term tests, such as measuring effects on photosynthesis, nutrient cycling, degradation of organic matter, energy conversion, survival etc, on a “community sample” that is challenged with known toxicants in the laboratory. Common endpoints are therefore e.g. thymidine incorporation into nucleic acids of bacteria, but also nematode lethality (Millward & Grant 1995, 2000), and abundance in benthic invertebrate assemblages (Courtney & Clements, 2000).

Thus, the approach works for any community that can be sampled, and PICT was so far used to assess the tolerance developed by nematode and other invertebrates, periphyton, phytoplankton and bacteria communities in both marine and limnic environments (Blanck 2002). PICT was primarily used for risk assessment purposes to assess the risks of individual contaminants on community levels (as opposed to the traditional approach to measure effects on single species) by exposing sampled communities collected from clean sites in the laboratory (Blanck 2002). PICT was also used in retroactive risk assessment studies on marine periphyton communities, by sampling communities from contamination gradients and exposing the samples to single chemicals known to be contained in these gradients, such as TBT and irgarol (Blanck & Dahl 1996; Blanck et al 2009). Recovery was also studied before and after the TBT ban (Blanck & Dahl 1998).

In recent years, PICT combined with the transplantation of periphyton communities has been suggested as a promising tool to identify impaired sites by detecting an induced tolerance after transplantation. Transplantation techniques of periphyton communities are facilitated by utilising the rapid colonisation occurring on deployed glass discs. In situ PICT assays using transplanted communities has been suggested as a promising tool that can link ecological and chemical status in the WFD context (Pesce et al 2010a, b; Tlili et al 2010, 2011).A disadvantage of the PICT-approach is that it cannot be used to assess the risks for long-lived organisms with complex life cycles (e.g. insects, vertebrates). One also needs to be aware of the possibilities that the organisms can develop co-tolerance to chemicals with a similar mode of action.

The toxicants suspected to change community function needed so far to be identified in order to know what substance or mixture that should be used in the short term test. In an environment that is influenced by complex mixtures from many sources, including substances that are rapidly degraded or transformed, it may be difficult to decide which single substances that are the most relevant to use at least in a monitoring context. However, a potential use within the WFD context would be to aid in the identification of specific pollutants based on ecologically relevant data. If a limited number of substances can be suspected to be of a major concern in a water body, but assessment criteria were not developed/uncertain, or the combined effects from these substances are hard to predict, the PICT approach offers a possibility to directly investigate effects on community

levels (on lower trophic levels) from these individual substances in the water body of concern. It can also distinguish which of the individual substances tested that are the most important from an ecological perspective, and should therefore be valuable decision support from a management perspective.

The PICT approach is also being developed and in combination with new tools, its usefulness in a WFD context could become broader. By exposing communities sampled from ”clean” environments to water or sediment samples from contaminated (downstream) environments, either in the laboratory or in the field, and measure effects on relevant functions, it would be possible to actually avoid the step of identifying the suspected contaminants before being able to measure the effects (Rotter et al 2011). The lack of knowledge on causing agents from the start, would necessitate the use of endpoints that can provide integrated response from several potential mode of actions. So far, tymidine incorporation in bacteria would be the only such identified endpoint used within PICT studies (Blanck 2002). However, Montuelle et al (2010) concluded that PICT combined with genetic fingerprints and different OMICS tools (see next section) offer additional possibilities.

6 EDA and TIE

While effect-based monitoring indicates hazards due to chemical contamination and provides information on toxicological endpoints of concern, tools are required to identify causes and to elucidate links between exposure and effect (e.g. intersex; Desbrow et al 1998). EDA and TIE are integrated biological and chemical approaches to identify those compounds in an environmental or technical sample (water, soil, sediment, air, food, consumer product, technical mixture) that cause a biological response. Both approaches combine biotesting, physico-chemical fractionation and chemical analysis in a sequential procedure. However, the philosophy behind both approaches is slightly different (Burgess et al., 2013).

The TIE approach has its origin in whole effluent testing, which focuses on the question, whether an effluent will cause adverse effects on aquatic organisms when emitted to the environment. In the case that effects are detected in whole organisms under realistic exposure conditions TIE should help to characterise and identify the cause of the measured effect. Thus, TIE applies *in vivo* biotesting and avoids extraction and pre-concentration steps as far as possible. Guidelines for TIE of water and sediment samples have been provided by US-EPA. The procedure involves three steps: (1) Toxicity characterisation is applied to link observed effects to groups of chemicals such as metals, lipophilic organic compounds, volatiles or ammonia using simple sample manipulations with subsequent biotesting. (2) Toxicity identification applies fractionation procedures and chemical analysis to identify candidate toxicants. (3) Toxicity confirmation is designed to confirm the identified toxicants as the cause of the measured effects. TIE has a been widely used in the U.S. and has been shown to be a powerful approach to characterise effluents as well as water and sediments from contaminated sites causing toxicity to aquatic organisms. The specific focus of TIE on *in vivo* toxicity under realistic exposure conditions is a major strength of this approach but also its main limitation. If no *in vivo* toxicity of the original sample can be detected – and this is the case in the vast majority of European surface waters - TIE is not applicable. However, the absence of acute toxicity in surface waters does not necessarily indicate that no chemicals effects on the community will occur. Several studies observed for example the disappearance of sensitive species at concentrations that are a factor of thousand below acute EC50 values (von der Ohe 2009) [22]. This highlights the requirement of sensitive detection of sublethal effects for toxicant identification that goes significantly beyond what is typically done in TIE. .

Effect-directed analysis (EDA) is based on the understanding that environmental samples may contain thousands of mostly organic chemicals and that only a fraction of them can be analysed by chemical target analysis. Non-target screenings provide valuable insights into unexpected or unknown contaminants but they do not provide any prioritisation or hazard information. Thus, EDA takes a biological effect (typically observed by effect-based monitoring) as the basis to narrow down the huge amount of possible analytes and aims to direct chemical analysis to those compounds that contribute significantly to a measurable effect. Thus, in EDA bioassays are considered as tools to sensitively detect chemicals with similar biological targets or modes of action. The focus of EDA is on unraveling the contamination with organic toxicants representing the most complex group of chemicals. Similar to chemical analysis there are no restrictions with respect to extraction or pre-concentration. Since the isolation and identification of individual toxicants out of thousand of components in typical environmental mixtures often

demand for large numbers of fractions high-throughput tools are preferred. In addition, the identification of unknown toxicants is very much supported by information on the mode of action. Both criteria are often met best with in vitro assays, although small scale in vivo assays may be helpful, too. The approach combines biotesting, physico-chemical fractionation procedures and chemical analysis in a sequential procedure (Figure 1, Brack 2003 [23]). The sample or an extract thereof is tested with the biotests of choice depending on the objective of the study. If effects are detectable the mixture is fractionated according to the physico-chemical properties of the components. The fractions are tested with the same biotests for prioritisation according to effects. The mixture may undergo several fractionation steps to further reduce complexity. The components of active fractions are identified and quantified by chemical analytical means. Depending on the objective of the study, in a final confirmation step the contribution of the identified candidate compound to the measured effect should be quantified or estimated in order to exclude that major contributors have been overlooked.

6.1 Components of EDA

The major components of EDA are (i) separation including extraction, clean up and fractionation, (ii) biotesting, (iii) chemical analysis including computer tools for structure elucidation, and (iv) confirmation.

(i) EDA typically starts with an extraction and pre-concentration step. If water samples need to be analysed passive sampling techniques or active solid phase extraction (SPE) may be used to remove the chemical mixture from water. Since EDA is targeted to identify unknowns typically extraction techniques are applied that pre-concentrate compounds with a broad spectrum of physico-chemical properties. Passive sampling typically extracts compounds over a time period of days or weeks and thus may be more representative than grab sampling with subsequent SPE. However, the identification and prioritisation of unknowns is restricted by unknown individual sampling rates and thus, the mixture found in the sampler may significantly differ from the mixture found in the water. To avoid this bias in most cases active sampling with SPE is preferred. For cost-efficient monitoring involving bioassays and EDA there is a need for mobile high-volume SPE methods that can be applied in situ. Care should be taken to apply SPE adsorbents for a broad range of compounds. This requirement can be met by using combinations or mixtures of different sorbents. Kern et al successfully applied for example a mixture of Strata-X-AW, Strata-X-CW (both Phenomenex), Isolute ENV+ (Separtis) and Oasis HLB (Waters) (Kern et al., 2009).

If sediment samples are in the focus of analysis EDA may be based on pore water or elutriates, exhaustive organic extracts or apply bioaccessibility-directed extraction tools (e.g. (Schwab and Brack, 2007). Although pore water and elutriates seem to reflect bioavailability in a better way than organic extracts they are rarely used in EDA. Reasons are the practical problems to gain sufficient amounts but also the potential underestimation of bioavailable contamination. While in sediments losses from the pore water by adsorption, degradation und uptake into organisms are compensated by desorption from sediments in the laboratory storage and test vessels free of sediments this compensation may not occur.

Before samples can be directed to biotesting and fractionation a clean-up step may be required. Since sediments like all environmental samples may contain toxicants with a great range of polarity which should be considered in EDA clean up often applies a separation according to the molecular size using size exclusion chromatography or dialysis

procedures (Streck 2008). These techniques remove large interfering compounds such as humic compounds, lipids or proteins, which are considered as matrix.

Fractionation is predominantly based on preparative reversed phase (RP) and normal phase high performance liquid chromatography (NP-HPLC). Automated multistep fractionation procedures have been developed to increase selectivity and throughput particularly in NP-HPLC (Lubcke-von Varel et al., 2008). Preparative capillary gas chromatography (pcGC) may be used for final isolation of toxicants (Meinert and Brack, 2010).

(ii) EDA can be based on any toxicological, ecotoxicological or biological endpoint that can be detected and quantified with sufficient throughput. Since there is no toxicity as such but interactions of chemicals with specific biological systems, biotest batteries covering different modes of action provide more comprehensive information than individual test systems. *In vivo* test batteries may for example test for toxicity to algae, invertebrates and bacteria to cover baseline toxicity but also neurotoxicity and effects on photosynthesis (Brack et al., 1999). *In vitro* assays applied in EDA include mutagenicity, tumor promotion and several types of endocrine disruption effects (Thomas et al., 2004b; Thomas et al., 2001).

Although many biological processes and functions can be disturbed by environmental chemicals the number of toxicological endpoints covered by available bioassays is rather limited. Thus, it will be an important task for the future to better cover the complexity of life and the complexity of interactions with chemicals and other stressors with bioassays and other bioanalytical tools. “Omics” techniques may help to identify important pathway of toxicity of environmental pollution and help to extend and improve bioassay batteries or maybe even serve as diagnostic tools in EDA themselves.

(iii) The chemical identification is still one of the major challenges in EDA. Gas chromatography with mass spectrometry (GC-MS) is by far the most frequently applied technique for toxicant identification. Extensive spectra libraries support identification and may be supplemented by structure generation tools and by computer tools and models providing fragmentation- and retention-based classifiers for structure elucidation. GC-MS techniques are limited to non- and medium polar compounds that can be evaporated without decay while polar and thermolabile compounds are of increasing importance in the environment. Thus, although less elaborated, LC-MS techniques play an increasingly prominent role in EDA ((Hogenboom et al., 2009); (Bataineh et al., 2010). Other techniques for structure elucidation such as NMR spectroscopy may be helpful in some cases (Nukaya et al., 1997). However, insufficient amounts and purity of toxicants in environmental samples often prevents their application.

(iv) Toxicant confirmation in EDA is required to provide evidence that the identified compounds actually explain at least part of the measured effects and are of relevance in the analyzed sample. Confirmation may be regarded as a tiered approach (Figure 2, Brack 2008) including analytical confirmation and effect confirmation with the respective biotest. Depending on the objective of EDA a final tier may be hazard confirmation under realistic exposure conditions and on a higher level of biological organisation.

6.2 Bioavailability in EDA

Hazards and risks due to soil and sediment contaminants depend on adverse effects and bioavailability. To reflect this in prioritisation of fractions and compounds in EDA several approaches are available including bioaccessibility-directed extraction techniques ((Schwab et al., 2009), passive dosing mimicking partitioning processes in sediments (Bandow et al., 2009) and EDA in body fluids or tissues of exposed organisms (Hewitt et al., 2003; Houtman et al., 2004).

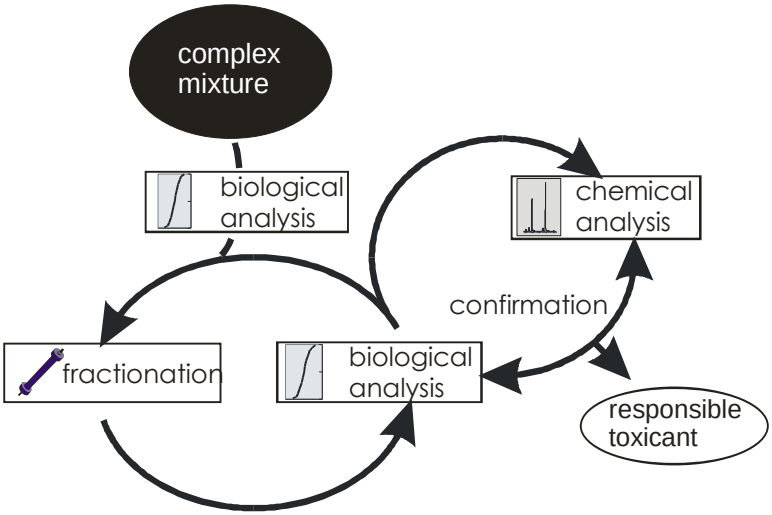


Figure 1: Scheme of effect-directed analysis of complex mixtures (Brack, 2003)

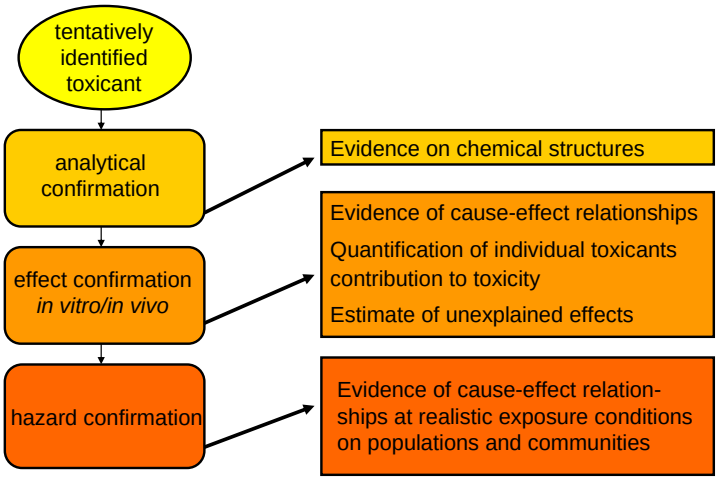


Figure 2: Toxicant confirmation as a tiered approach (Brack et al., 2008)

6.3 Some success stories

Endocrine disruption is a major concern related to environmental pollution with significant relevance for human health but also for aquatic ecosystems. In extensive EDA studies the natural steroids 17β-estradiol and estrone as well as the synthetic hormone 17α-ethynylestradiol were identified as major estrogens in UK rivers and estuaries. All of these compounds stem from domestic sewage treatment works (STP). The same natural and synthetic steroids were identified by (Houtman et al., 2004) as major estrogens when applying EDA to bile of breams in River Dommel, The Netherlands. Other groups were able to link estrogenic effects to nonylphenol and related chemicals (Cespedes et al., 2004). Androgenicity could be explained by natural steroids and their metabolites including dehydrotestosterone, androstenedione, androstanedione, 5β-androstane-3α,11β-diol-17-one, androsterone and epi-androsterone.

In sediments of European rivers several androgen-disrupting chemicals have been identified by EDA including steroids such as androstenone and nandrolone but also xenobiotics including the musk compound galaxolide and tris-(2-chloroisopropyl)phosphate (Weiss et al., 2011).

The case study “Contaminated sediments in the River Elbe basin” describes several additional sediment contamintants of concern that could be identified, when using a large bioassay battery.

Mutagenicity was frequently detected in Japanese river waters and considered as a concern for human health and ecosystems. Applying large amounts of blue rayon as passive samplers designed to adsorb polycyclic aromatic compounds of different polarities, the Ames test with different diagnostic strains, extensive fractionation procedures and high resolution MS and ^{1}H NMR Japanese groups were able to identify several 2-phenylbenzotriazole type mutagens from azo dye production as the cause of mutagenicity (Shiozawa et al., 2000).

6.4 EDA and WFD

While effects monitoring based on a broad array of toxicological endpoints is available for surveillance and operational monitoring, EDA is a tool for investigative monitoring at selected sites of particular interest or with conspicuous effects. EDA helps to link ecological status with contamination, to establish cause-effect relationships and to target mitigation measures. To further broaden its applicability simplified protocols and high-throughput approaches need to be developed and tested in the field.

Although providing enormous progress over present target chemical monitoring a general limitation of effect-based monitoring and EDA is the requirement to pre-select toxicological endpoints. The combination of integrating whole organism tests with *in vitro* test batteries applying sufficient pre-concentration reduces the risk to overlook important effects and thus toxicants. It may be expected that emerging “omics” techniques may reduce the limitations due to endpoint-specificity by allowing analysis of a holistic health status and many endpoints at the same time. Major progress may be also expected with respect to the isolation and identification particularly of polar compounds with LC and LC-MS techniques. Technical progress in instrumentation, the application of emerging chromatographic techniques and stationary phases, the development of libraries and databases, the advancement and application of computer tools to predict fragmentation

and retention and the combination with QSAR (Thomas et al., 2004a) techniques for effect prediction will significantly enhance the analytical power of EDA and will help to extend it to new matrices and problems

7 OMICS Technologies

7.1 Introduction

The recent advances in sequencing and characterisation of genomes have opened up new possibilities. A particular field of molecular studies within biology is called Omics and refers to high throughput molecular profiling technologies, such as Genomics, Metagenomics, Proteomics or Metabolomics. The suffix “-ome” refers to the collection of all genes or gene products such as the genome, proteome or metabolome respectively. So a study of all or a very large number of these genes would fall under the definition of omics. Other types of molecular analyses that can also be valuable in this context (but generally analyse much less numbers of gene products at the same time) are e.g. qPCR (being described in more detail in the Annex, section 9), Western Blot and ELISA.

Whereas Genomics is the study of the genomes of organisms, *Meta*genomics is the study of the genetic material of the whole community. Transcriptomics is the study of the transcriptome which is the set of all RNA molecules, including mRNA, rRNA, tRNA, and other non-coding RNA produced in one cell or a population of cells. In proteomics the proteome, which is the entire complement of proteins, including the modifications made to a particular set of proteins, produced by an organism or system, is studied. In metabolomics the chemical processes involving metabolites is studied, such as the unique chemical fingerprints that specific cellular processes leave behind.

Omics and bioinformatics tools can e.g. be used to

- develop molecular biomarkers of exposure as early signals to predict effects (that at a later stage could have an impact on physiological level, and later on at population level).
- provide information about the mode of action (MOA) of chemicals, i.e. the mechanism of toxicity; in turn reducing the uncertainties involved in chemical risk assessment by providing, for example, a basis for the extrapolation of the effects across species.
- integrate the data of MOA with a deleterious outcome, and in this way understand the impact on the ecosystem more than only on single organism or species.
- distinguish the site of origin of organisms, based on the transcriptomics changes in organisms coming from different locations (see Annex, section 10).

7.2 Genomics-DNA Microarray applications

Genomics can tell you about the susceptibility of an organism for a certain chemical (see e.g. Gunnarsson et al 2012). A DNA Microarray is a glass or a nylon membrane on which part of gene sequences (probes) are spotted. Normally, complementary DNA (cDNA) is made using reverse transcriptase from RNA. Then the cDNA is hybridised to the chip. After scanning image analysis, the RNA abundance (amount of RNA molecules bound to the

complementary probes on the microarray) is analysed and therefore the relative gene expression of the treated sample can be compared to the untreated control.

A major benefit of using the DNA Microarray technique is the possibility to investigate in one experiment the response of an entire organism to a pollutant, when the array for the species´ genome is available. It is also possible to target clusters of genes linked to specific organs such as the ovary (Larkin et al., 2007) and to analyse that specific gene expression upon exposure to chemicals.

Today, microarray analyses are established analytical tools in many laboratories and used for several applications. Applications of genomic experiments in aquatic toxicology have been described in several scientific papers (see Overview of existing DNA microarrays in Annex, section 11)..

Previously, the high cost of microarrays imposed several restrictions in terms of the number of biological experiments, time points, and concentrations that were possible to analyse. Consequently, the experimental approach mainly focused on determining an EC_{50} and not NOEC or testing environmentally relevant concentrations. Nowadays, the reduced price has allowed the improvement of the methodology as well as the possibility to screen for several chemicals at the same time. In addition, many commercially or customised DNA microarrays are available for many species such as diatoms (Carvalho et al*.,*2011a), the planktonic crustacean *Daphnia magna* (Watanabe et al., 2007) and several fish species, such as zebrafish (Mathavan et al., 2005) and fathead minnow (Wintz et al., 2006). The Annex (section 11) provides an overview of existing DNA microarrays for several organisms and the stressors tested as well as type of environmental samples.

The availability of these DNA microarrays increases the number of non-model organisms, providing more data on the dose-response relationship of chemicals, but also the differential species sensitivity, and the classification of chemical-specific biological responses (Van Aggelen et al., 2010). What is now lacking is a standardised approach, where gene/protein/metabolite expression profiles are combined with chemical exposure to develop a toxicity-profile for environmental monitoring. It would facilitate a more accurate and reliable analysis of DNA microarrays, allowing the comparison of mode of action of several pollutants between test species. These data will identify specific and common gene signatures to be used for molecular based bioassay developments, such as quantitative real Time PCR screening for environmental monitoring, see 5.1.2. The case study “Evaluation of the utility of microarrays as a biomonitoring tool in field study” also illustrates how microarrays were used to distinguish between more and less heavily impacted sites exposed to complex chemical mixtures, and the case study “Use of DNA microarray to test the water quality of river East Turkey Creek (bay of watershed of Florida) potentially impacted by treated wastewater from sprayfield area” illustrated how these tools can be used to provide insight into the water quality degradation (Annex, section 1). The Annex (section 10) also includes some additional examples on the development of molecular biomarkers and how the site of origin could be tracked using an OMICS approach.

7.3 Next-generation sequencing (NGS)

The development of DNA sequencing technology four decades ago was a major scientific hallmark and opened the doors for inumerous breakthough achievements in all areas of

biology. Next-generation sequencing (NGS) is a more recent technology, also named second-generation sequencing (SGS) and has been commercially available since 2004. Compared to the first-generation capillary electrophoresis (CE)-based Sanger sequencing, which could process 96 samples at a time, NGS has increased the throughput to millions of sequencing experiments on fragmented DNA run in parallel. NGS platforms enable a wide variety of applications including the study of the genome or transcriptome of any organism. There are a few NGS platforms on the market, differing in the amplification approach, nucleotide determination method, time per run and read length. In addition, recent development of single molecule sequencers, so-called "third-generation sequencers", allows the sequencing of samples without amplification by means of True Single Molecule Sequencing (tSMS) technology, with the potential for longer read lengths, shorter time and lower overall cost.

7.3.1 RNA-seq

RNA-seq is a recently developed approach, extending the high-throughput sequencing to the profiling of the transcriptome. Instead of capturing transcript molecules by molecular hybridisation, like on a microarray, RNA-Seq directly sequences the transcripts present in a sample. Transcript sequences are then mapped back to a reference genome and counted to assess the expression level of that gene or genomic region (Wang et al. 2009). Thus, it provides a direct quantification of gene expression instead of a proxy measurement of RNA abundance (fluorescence as a function of hybridisation) inherent to molecular hybridisation technologies. RNA-seq is particularly useful for de novo investigation, studies on alternative splicing or on non-coding RNA. RNA-seq has many advantages that include low background, high sensitivity, high throughput and low amount of RNA required. Unlike hybridisation-based approaches, RNA-Seq is not limited to detecting transcripts that correspond to known genomic sequences, which is particularly useful for non-model organisms with unknown genomic sequence. As an example, RNA-seq was applied recently to the non-model but ecologically relevant organism the salt marsh minnow *Fundulus grandis*. The study assessed the genomic changes in *F. grandis* populations in habitats impacted by oil release following the blowout of the Deepwater Horizon drilling platform in the Gulf of Mexico (Garcia et al, 2012).
However, RNA-seq still poses a few challenges, the biggest of all in the analysis of the imense amount of data generated, including the development of bioinformatics tools, and the cost. While some will argue RNA-seq will outplace hybridisation-based gene expression analysis in a few years, it is more likely that RNA-seq will complement and extend microarray experiments, with the choice depending largely on the research project.

7.3.2 Metagenomics – A novel tool to study ecosystem function

In recent years a tremendous increase in DNA sequencing capacity, combined with an unprecedented drop in price per obtained nucleotide sequence, have made it possible to study the functional elements of an ecosystem at the levels of the actual genes responsible for these functions. Such sequencing studies of the total DNA content of an environment is generally referred to as metagenomics (Riesenfeld et al. 2004).

7.3.2.1 Total microorganism community function

Metagenomic studies can be targeted towards specific genes of interest, such as taxonomic markers (16S rRNA) or genes involved in detoxification or antibiotic resistance development. However, it is often of interest to study the total functional content of a community of microorganisms to get a broader picture of the ecosystem function (Tringe et al. 2005). To assess such broad questions about the total composition of organisms, genes and functions represented within a community in a single experiment would have been impossible without the leap forward in sequencing technology seen in the last decade (Metzker 2010). Metagenomics provides a means to analyse complete communities of microorganisms, regardless of whether they can be cultured in the laboratory or not. This is a huge benefit as it has been estimated that only one or a few per cent of the microorganisms in nature can be readily grown in the laboratory (Amann et al. 1995).

7.3.2.2 Standardisation potential and reproducibility

One of the main benefits of using DNA sequencing and metagenomics for community analysis is that there are easily implemented and standardised protocols for DNA extraction and amplification, and that the methods produce reliable and reproducible results. There are several laboratories and companies that include most of the DNA preparation as part of the sequencing service, making the process from extracted DNA sample to resulting sequences highly standardised and more or less transparent to the end-user.

7.3.2.3 Data interpretation and computational capacity to decide functional role of each sequence

Because of the large amount of sequences generated in a single sequencing run, most often on the order of hundreds of thousands to hundreds of million sequences, the major challenge of metagenomics is the post-sequencing analysis in which the functional role of each sequence is determined. This means that there is a great need for bioinformaticians to make sense of the DNA information. Currently, there is no easy-to-use software solution that can harness the power of metagenomics without the need for some degree of bioinformatics expertise. Such solutions are, however, undoubtedly in development, and progress has been made with packages such as MEGAN (Huson et al. 2011). Nevertheless, either in-house bioinformaticians or consultants are today a requirement to be able to draw conclusions from the vast amount of sequence data generated using the modern sequencing techniques. In addition, the large amount of DNA sequenced in these studies requires substantial storage and computing capacities.

7.3.2.4 Measurement of species composition

An important benefit of using metagenomics for community studies is that the amount of sequences generated enables very precise measurements of e.g. the species composition in a community. Based on a genetic marker approach it is possible to study the species composition in the community, in terms of which species or groups of species that are present and how different toxicant selection pressures affect the community structure. Previous knowledge of the exact sequence of the genetic marker in a given species is not needed to be able to identify it and give an estimate of its taxonomic association.

7.3.2.5 Ecological relevance and possibilities to identify previous exposure to toxicants

The study of species composition changes in a community is not a sub-lethal endpoint but rather an analysis of effects on the community and ecosystem levels. The ability to investigate highly and lowly represented genetic functions in a community translates into an understanding of which functions and organisms that are required to maintain ecosystem stability, and in the end the sustainability of the ecosystem services] provided by the community.

The ability to study functions present or absent in a community, as well as the abundance of the genes responsible for these functions, also makes it possible to assert something about previous exposure to various toxicants. This assessment is either based on sensitive versus non-sensitive indicator species found, or by looking at the abundance of genes encoding specific detoxification functions, cf SPEAR and PICT. It would also be possible to e.g. combine studies of community tolerance, such as PICT, with metagenomics to investigate the underlying changes in genetic composition responsible for inducing a response to the toxicant at the phenotypic level.

A major benefit of using metagenomics to study the composition of microbial communities is the possibility to not only focus on specific genes and functions, but to look broader into *all* available genes and functions. In practice, this means that it is not necessary to propose a hypothesis before conducting the actual experiments. Instead, the content of organisms and functions detected in the sample can be used to trace effects of toxicants, regardless of whether the presence of any specific toxicant was known on beforehand.

7.3.2.6 Antibiotic resistance genes in the environment

Some antibiotics were considered for listing as priority substances during the most recent review of the list on the basis of their toxic effects. However, another major concern related to antibiotics in the environment (especially if released from sewage treatment plants where there are biological treatment steps with recirculation) is the potential release of antibiotic resistance genes, that over time could end up in pathogenic bacteria and pose a health hazard. A possible use of metagenomics is to monitor the presence and abundance of antibiotic resistance genes in the environment. While this has been possible for a long time using culturing techniques, metagenomics provides a means to look into the resistance profile of *all* bacteria in a community. This is a benefit because of the limited number of cultivable microorganisms in nature, as pointed out earlier.

7.4 Proteomics

Proteomics is the large-scale study of proteins, particularly their structures and functions Andersson and Andersson, 1998; Blackstock and Weir 1996). After genomics and transcriptomics, proteomics is considered the next step in the study of biological systems. It is much more complicated than genomics mostly because while an organism's genome is more or less constant, the proteome differs from cell to cell and from time to time. This is because distinct genes are expressed in distinct cell types. This means that even the basic set of proteins which are produced in a cell needs to be determined. Proteomics typically gives us a better understanding of an organism than genomics. First, the level of transcription of a gene gives only a rough estimate of its *level of expression* into a protein

(Gygi et al 1999). An mRNA produced in abundance may be degraded rapidly or translated inefficiently, resulting in a small amount of protein. Second, as mentioned above many proteins experience *post-translational modifications* that profoundly affect their activities; for example some proteins are not active until they become phosphorylated. Third, many transcripts give rise to more than one protein, through alternative splicing or alternative post-translational modifications. Fourth, many proteins form complexes with other proteins or RNA molecules, and only function in the presence of these other molecules. Finally, protein degradation rate plays an important role in protein content (Belle et al, 2006).

Proteomics experiments conducted in one laboratory are not easily reproduced in another. For instance, Peng et al. (2003) have identified 1504 yeast proteins in a proteomics experiment of which only 858 were found in a similar previous study (Washburn et al, 2001). Further, the previous study identified 607 proteins that were not found by Peng et al (2003). This translates to a reproducibility of 57% (Peng vs. Washburn) to 59% (Washburn vs. Peng).

7.5 Metabolomics

Within metabolomics the endogenic metabolic profile of an organism is studied. The metabolites that are studied can be considered to be the result of the ongoing metabolic activity of the cells. To measure metabolites is considered advantegous since it is well known that metabolites are formed at an early stage of environmental stress (Sternbeck et al, 2008). However, because the results possibly are rather complex, the measurements of metabolites need to be combined with a multivariate analysis in order to be able to establish dose response relationships. In an investigation performed by Samuelsson (2006) metabolomics was used to study the impact of phamaceuphticals and other xenobiotics in the aquatic environment. It was shown that effects were found at lower concentrations of EE2 with metabolomics compared to measurements of vitellogenin.

KH-04-14-117-EN-N

Publications Office