CLH report

Proposal for Harmonised Classification and Labelling

Based on Regulation (EC) No 1272/2008 (CLP Regulation),
Annex VI, Part 2

Substance Name:

2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol)

EC Number: 403-800-1
CAS Number: 103597-45-1
Index Number: 604-052-00-0

Contact details for dossier submitter:

BAuA
Federal Institute for Occupational Safety and Health
Federal Office for Chemicals
Friedrich-Henkel-Weg 1-25
D-44149 Dortmund, Germany

Dossier prepared by:
Industry in accordance with Article 37(6) of CLP Regulation

Version number: **2.0**
Date: **September 2016**

CONTENTS

Part A.

1 PROPOSAL FOR HARMONISED CLASSIFICATION AND LABELLING

1.1 Substance

Table 1: Substance identity

Substance name:	2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol)
EC number:	403-800-1
CAS number:	103597-45-1
Annex VI Index number:	604-052-00-0
Degree of purity:	100 %
Impurities:	*Impurities are considered to be confidential to the public*

1.2 Harmonised classification and labelling proposal

Table 2: The current Annex VI entry and the proposed harmonised classification

	CLP Regulation
Current entry in Annex VI, CLP Regulation	Aquatic Chronic 4
Current proposal for consideration by RAC	Removal: Aquatic Chronic 4
Resulting harmonised classification (future entry in Annex VI, CLP Regulation)	None

1.3 Proposed harmonised classification and labelling based on CLP Regulation

Table 3: Proposed classification according to the CLP Regulation

CLP Annex I ref	Hazard class	Proposed classification	Proposed SCLs and/or M-factors	Current classification [1)]	Reason for no classification [2)]
4.1.	Hazardous to the aquatic environment	None		Aquatic Chronic 4	Conclusive but not sufficient for classification

[1)] Including specific concentration limits (SCLs) and M-factors
[2)] Data lacking, inconclusive, or conclusive but not sufficient for classification

Labelling: Signal word: no signal word
Hazard statements: no H-statements
Precautionary statements: no precautionary statements

Proposed notes assigned to an entry: none

2 BACKGROUND TO THE CLH PROPOSAL

The dossier was prepared by industry according to Article 37(6) of CLP Regulation.

For the purpose of this dossier the German CA has taken all registration dossiers available in September 2016 into account. Nevertheless, not all available studies for aquatic toxicity were listed in this dossier since all studies show the same results (no effects in the range of the water solubility).

2.1 History of the previous classification and labelling

The harmonised classification (R 53) of 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) had been included in 67/548/EEC with the 26th ATP.

According to EC/1272/2008 Annex VI, the substance may cause long lasting harmful effects to aquatic life and thus, meets the criteria for classification with Aquatic Chronic 4. This classification is based on the high logPow value (> 3), the resulting bioaccumulation potential of the substance, non rapid biodegradability, no acute toxicity up to the water solubility and the absence of chronic toxicity data on both aquatic invertebrates and fish.

2.2 Short summary of the scientific justification for the CLH proposal

New experimental data show that 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) has no chronic effects towards algae and aquatic invertebrates. According to the acute aquatic toxicity data, neither fish nor aquatic invertebrates seem to be more sensitive. A chronic fish toxicity test is therefore not necessary to assess the toxicity towards aquatic organisms. Furthermore, the bioaccumulation potential is expected to be low based on the available information from BCF QSAR calculations, mammalian toxicokinetic studies, logPow and water solubility. Therefore, classification of the substance with Aquatic Chronic 4 is no longer justified.

2.3 Current harmonised classification and labelling

2.3.1 Current classification and labelling in Annex VI, Table 3.1 in the CLP Regulation

Table 4: Current classification and labelling in Annex VI, Table 3.1 in the CLP Regulation (Index-No.: 604-052-00-0)

Classification		Labelling			Specific Conc. Limits, M-factors	Notes
Hazard Class and Category Code(s)	Hazard Statement Code(s)	Hazard statement Code(s)	Suppl. Hazard statement Code(s)	Pictogram, Signal Word Code(s)		
Aquatic Chronic 4	H413	H413				

2.4 Current self-classification and labelling:

The following industry self-classification(s) and labelling are publically available in the ECHA C&L Inventory.

Table 5: Current industry self-classifications(s) and labelling in the ECHA C&L Inventory (September 2016)

Classification		Labelling		Specific Concentration limits, M-Factors	Notes	Number of Notifiers
Hazard Class and Category Code(s)	Hazard Statement Code(s)	Hazard Statement Code(s)	Pictograms, Signal Word Code(s)			
Aquatic Chronic 4	H413	H413				65 (joint entry)
Not classified						3 (joint entry)
Aquatic Chronic 4	H413	H413				75

3 JUSTIFICATION THAT ACTION IS NEEDED AT COMMUNITY LEVEL

According to new data, modification of the existing entry is appropriate. The classification and labelling as Aquatic Chronic 4 is not justified.

Part B.

SCIENTIFIC EVALUATION OF THE DATA

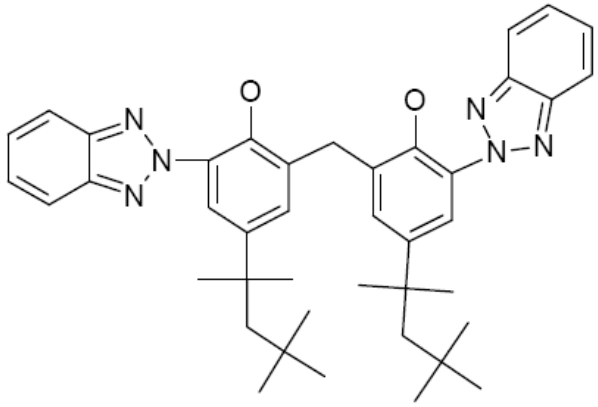
1 IDENTITY OF THE SUBSTANCE

1.1 Name and other identifiers of the substance

Table 6: Substance identity

EC number:	403-800-1
EC name:	2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol)
CAS number:	103597-45-1
CAS name:	Phenol, 2,2'-methylenebis[6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)-
IUPAC name:	2-(benzotriazol-2-yl)-6-[[3-(benzotriazol-2-yl)-2-hydroxy-5-(2,4,4-trimethylpentan-2-yl)phenyl]methyl]-4-(2,4,4-trimethylpentan-2-yl)phenol
CLP Annex VI Index number:	604-052-00-0
Molecular formula:	$C_{41}H_{50}N_6O_2$
Molecular weight:	658.89 g/mol

Structural formula:



1.2 Composition of the substance

Table 7: Constituents (non-confidential information)

Constituent	Typical concentration	Concentration range	Remarks
2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol)	99.3 % (w/w)	95.0 – 99.9 % (w/w)	

Current Annex VI entry: Aquatic Chronic 4; H413

Table 8: Additives (non-confidential information)

Additive	Function	Typical concentration	Concentration range	Remarks
none				

Impurities are considered to be confidential and are stated in the technical dossier.

1.2.1 Composition of test material

The test material is a mono-constituent substance.

1.3 Physico-chemical properties

Table 9: Summary of physico-chemical properties

Property	Value	Reference	Comment (e.g. measured or estimated)
State of the substance at 20 °C and 101.3 kPa	Solid (powder)	Ciba- Geigy LTD (1991)	
Melting/freezing point	195.7 °C	Ciba- Geigy LTD (1991)	
Boiling point	571.7 °C at 1013 hPa (extrapolated) 276.2 °C at 11 Pa	Ciba- Geigy LTD (1991)	
Relative density	1200 kg/m³ at 22 °C	Ciba- Geigy LTD (1991)	
Vapour pressure	0.000000000006 Pa at 25 °C	Ciba- Geigy LTD (1991)	
Surface tension	not applicable	Expert judgement	The water solubility is < 1 mg/l
Water solubility	<0.000005 mg/L at 20 °C	Ciba- Geigy LTD (1991)	
Partition coefficient n-octanol/water	12.7 at 25 °C (calculated)	Ciba- Geigy LTD (1991)	
Flash point	Not relevant	Expert judgement	Substance is a solid
Flammability	- Not highly flammable upon ignition - The substance has no pyrophoric properties and does not liberate flammable gases on contact with water.	Ciba- Geigy LTD (1991)	
Explosive properties	Not explosive	Ciba- Geigy LTD (1991)	
Self-ignition temperature	no self-ignition	Ciba- Geigy LTD (1991)	
Oxidising properties	non-oxidising	Ciba- Geigy LTD (1991)	
Granulometry	05% w/w= <40 µm 10% w/w= <63 µm 15% w/w= <100 µm	Ciba- Geigy LTD (1991)	
Stability in organic solvents and identity of relevant degradation products	is not considered to be critical	Expert judgement	
Dissociation constant	pKa = 7 at 25 °C (calculated)	Ciba- Geigy LTD (1991)	
Viscosity	Not relevant	Expert judgement	Substance is a solid

2 MANUFACTURE AND USES

Not relevant for the purpose of this dossier.

3 CLASSIFICATION FOR PHYSICO-CHEMICAL PROPERTIES

Not classified for physico-chemical properties.

4 HUMAN HEALTH HAZARD ASSESSMENT

Based on the available toxicological data, the substance 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) is not to be classified for human health hazard according to the criteria laid down in 67/548/EEC and regulation (EU) 1272/2008. The information given in this chapter is included as supportive information for discussions provided in Chapter 5.3.1, however, there is no intention for harmonization of toxicological endpoints. Besides the information given below, other toxicological data available are considered as not relevant for this dossier.

4.1 Toxicokinetics (absorption, metabolism, distribution and elimination)

In a toxicokinetic study in Wistar derived Alpk:AP_fSD rats according to OECD TG 417/427 and GLP, 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) has been applied topically 10% and 0.2% in the commercial cosmetic formulation or orally as a single dose of 50 mg/ kg bw.

For dermal treatment, the formulation comprised unlabelled and ^{14}C-radiolabelled test item homogeneously dispersed in the vehicle (Plantacare 2000, Xanthan gum, Propylene glycol and water) such that a dose of a set volume (100 µL/rat) was equivalent to the nominal dose level of 0.2 or 10 mg/rat. In each case, unlabelled test item (purity: 99.6 %) and ^{14}C-radiolabelled test item (radiochemical purity: 99.1 %) were mixed and milled to a particle size comparable to that of the commercial formulation, nominally 200 nm. The particle size of the milled test substance was determined by scanning electron microscopy (SEM) to be in the range 300 and 2000 nm, with a typical particle size of approximately 1000 nm. A single application of the formulated active ingredient to 10 cm² of skin was performed in 32 male rats. After dosing, the application sites were protected, but not occluded, using O-rings incorporating a nylon gauze cover. A strip of non-occlusive elasticized bandage was wrapped around the rat and over the application devices to help to hold them in place. Rats were housed individually in metabolism cages for the collection of urine and faeces. After a 6-hour exposure, the first two groups were terminated and the application sites of all the remaining rats were washed to remove the unabsorbed dose. Urine, faeces and cage wash were collected from each cage after the 6-hour skin wash, and then at daily intervals after dosing for the duration of each experiment. Groups of 4 rats were terminated at 6, 24, 72 and 120 hours after dosing. Under anaesthesia, the skin was washed to remove unabsorbed residual test item before exsanguination. The application site skin was then tape-stripped to remove the *stratum corneum*. The dose formulations and all samples, including selected tissues and residual carcasses were analyzed for radioactivity by means of liquid scintillation counting. Disintegration per minute (dpm) values were calculated using the appropriate quench correction data.
For assessment of the metabolic fate of 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) after oral application, 4 rats per sex were given a single oral dose of 50 mg/kg bw [^{14}C]-labeled test substance (radiochemical purity: 99.1 %). The excretion of radioactivity in urine and faeces was monitored via metabolism cages for 3 days after dosing. After this period, the rats were killed and residual radioactivity was measured in blood, selected tissues and the remaining carcasses. An additional group of 9 rats per sex received a single oral dose of 50 mg/kg bw [^{14}C]- labeled test substance and radioactivity was measured in blood and plasma over a 24-hour time course after dosing. Radioactivity in the samples was determined by liquid

scintillation counting. Analysis of metabolites was performed by HPLC – MS (Ion trap mass spectrometer). The dose formulations comprised unlabelled and radiolabelled test item suspended in 0.5 % (w/v) CMC in 0.1 % (w/v) aqueous Tween 80. Dose formulations were analysed for radioactivity content by liquid scintillation counting. The particle size of the milled test item was determined by scanning electron microscopy (SEM) to be in the range of 300 to 2000 nm.

Results for single dermal administration of 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol).

The homogeneity of the radiolabelled test item in both dose formulations was satisfactory throughout the periods of dosing. The test item was stable in both dose formulations for longer than their period of use in the study.

Following dermal exposure to the 0.2 % formulation for 6 hours, approximately 97 % of the applied radioactivity was removed from the skin surface by aqueous washing. Approximately 0.7 % (0.4 % was found in the *stratum corneum*) of the dose remained associated with the application site and some of this was available for absorption. However, the area under the curve (AUC) could not be calculated because of the non-detectable radiolabel in the blood. The residue associated with the application site remained low, and declined at later timepoints. The amount of dose absorbed remained similar at 0.2 - 0.8 % after 6, 24, 72 and 120 hours.

Following dermal exposure to the 10 % formulation for 6 hours, approximately 98 % of the applied radioactivity was washed from the skin surface. Approximately 0.2 % (0.1 % was found in the *stratum corneum*) of the dose remained associated with the application site following the 6-hour skin-wash and some of this was available for absorption. The residue associated with the application site remained similar at later time-points. The amount of dose absorbed remained similar at 0.2 - 0.4 % after 6, 24, 72 and 120 hours.

Table 10: Percutaneous penetration of 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) through rat skin in vivo over a 5-day period

	Time after application			
	6 hours	**24 hours**	**72 hours**	**120 hours**
Recovery of applied dose for the 0.2 % formulation (% or % ± SD)				
Total absorbed dose[a]	< 0.34	0.80 ± 1.20	0.27 ± 0.05	< 0.53
Total non-absorbed dose[b]	97.98 ± 1.59	98.62 ± 3.12	97.70 ± 1.95	99.07 ± 2.38
Total recovery	98.32 ± 1.72	99.42 ± 2.02	97.97 ± 1.99	99.60 ± 2.29
Recovery of applied dose for the 10 % formulation (% or % ± SD)				
Total absorbed dose[a]	< 0.21	< 0.41	< 0.18	0.34 ± 0.17
Total non-absorbed dose[b]	97.63 ± 4.63	98.06 ± 4.25	99.86 ± 4.24	101.08 ± 0.63
Total recovery	97.84 ± 4.68	98.46 ± 3.77	100.03 ± 4.22	101.42 ± 0.55

a: Sum of radioactivity recovered in urine, faeces, cage wash, bandage, tissues, GI tract with contents and carcass; given as percentage of applied dose

b: Sum of radioactivity recovered in 6-hour skin wash and/or terminal skin wash and *stratum corneum*, skin application site, covers and O-rings; given as percentage of applied dose

SD: Standard deviation of the mean value for 3 or 4 animals

Results for single oral administration of 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol).

Following a single oral dose of 50 mg/kg bw [^{14}C]- labeled test substance, excretion was rapid and extensive in male and female rats. Urinary excretion accounted for a mean total of < 0.01 % of the

dose for both males and females and faecal excretion accounted for mean totals of 96 and 97 % for males and females, respectively. Only one component, identified as the parent test substance, was found in the faecal extracts. Residues in tissues were very low (< 0.01 % of the dose). The radioactivity remaining in the residual carcass accounted for < 0.07 % of the dose for males and < 0.08 % for females. The concentration of radioactivity in blood and plasma was below the limit of detection at all time points up to 24 hours after dosing and the area under the curve (AUC) could thus not be calculated. The achieved mass balance was acceptable.
Based on analytical results, the mean achieved dose was 50.4 mg/kg bw, which was 101 % of the intended dose of 50 mg/kg bw.

Table 11: Recovery of administered radioactivity following single oral gavage application of 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3 tetramethylbutyl)phenol)

Excreta /Tissues	**% - Recovery**	
	Male (mean or mean ± SD)	**Female (mean or mean ± SD)**
Urine	< 0.01 ± < 0.01	< 0.01 ± < 0.01
Faeces	96.40 ± 2.63	96.90 ± 3.98
Cage wash	< 0.01	< 0.02
GI tract with contents	< 0.01	0.06 ± 0.05
Tissues and carcass	< 0.08	< 0.08
Total	96.48 ± 2.63	97.06 ± 4.00

SD: Standard deviation of mean values from 4 animals
GI: Gastro-intestinal

Conclusion

Following a 6-hour topical exposure, the *in vivo* dermal absorption of 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) from a 0.2 and a 10 % formulation was very low and accounted for not more than 0.8 % and 0.4 % of the dose, respectively over 5 days. The topically applied 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) did not achieve systemically measurable concentrations and was thus not bioavailable.
Under the conditions of this study, systemic availability of 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) was negligible after oral administration. The test substance 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) was quantitatively and rapidly excreted as parent compound via the faeces.

5 ENVIRONMENTAL HAZARD ASSESSMENT

5.1 DEGRADATION

Table 12: Summary of relevant information on degradation

Method	Results	Remarks	Reference
Directive 92/ CEE C	Half-lives estimated at 25°C: DT50 (pH = 4) = 488 hours DT50 (pH = 7) = 120 days DT50 (pH = 9) > 1 year	4 (not assignable)	ECHA CHEM (2015)
EEC, L 251 Vol. 27 (comparable to OECD 301B)	0 – 10 % CO_2 evolution after 28 d	1 (reliable without restrictions)	CIBA-GEIGY Ltd. (1991c)
84/499/EEC C.5 (comparable to OECD 301 B)	2% CO_2 evolution after 28 days	4 (not assignable)	ECHA CHEM (2015)
OECD 302C	0 % O_2 consumption after 28 d		RCC Ltd. (2005)

5.1.1 Stability

The substance is not expected to hydrolyze in water at environmental relevant conditions.

5.1.2 Biodegradation

5.1.2.1 Screening tests

Two studies on the ready biodegradability of 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) are available. The studies were conducted according to OECD guidelines 301B. Results show that the substance is not readily biodegradable in water (0 – 10% biodegradation after 28 days) (Ciba Geigy, 1991c; ECHA CHEM, 2015). These results were confirmed by an inherent biodegradability study (OECD 302C, 0 % biodegradation after 28 days) (RCC Ltd, 2005).

5.1.2.2 Simulation tests

No data available.

5.1.3 Summary and discussion of degradation

The substance is not rapidly degradable.

5.2 Environmental distribution

5.2.1 Adsorption/Desorption

Based upon a log Koc of 5.63 (adsorption/desorption screening test (soil, HPLC-method)), the substance has a high potential to adsorb on soil and sewage sludge (ECHA CHEM, 2015).

5.2.2 Volatilisation

Not relevant for this dossier.

5.2.3 Distribution modelling

Not relevant for this dossier.

5.3 Aquatic Bioaccumulation

Table 13: Summary of relevant information on aquatic bioaccumulation

Method	Results	Remarks	Reference
Cyprinus carpio aqueous (freshwater) flow-through Total uptake duration: 8 wk Method for Testing the Degree of Accumulation of Chemical Substances in Fish, MITI, July 13, 1974.	BCF: 0.1 — 1.5 (whole body w.w.) (Time of plateau: 2 wk) (steady state) BCF: <= 1.4 (whole body w.w.) (Time of plateau: 2 wk) (steady state) Lipid content: 4.2 % (start of exposure) (Weight, length and lipid content at the Initiation of exposure: weight average 23.1 g; length average 9.4 cm and lipid content average 4.2 %)	3 (not reliable) weight of evidence experimental result **Test material (EC name): 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol)**	Kyushu Chemical Biotesting Center (1986)

Table 14: Summary of relevant information on aquatic bioaccumulation: Predicted BCF values for applied QSAR models sorted by BCF (AD = Applicability Domain)

Model	BCF	In AD	Restraints	Reference
BCFBAF v3.01 (EPI Suite v4.11): **Arnot-Gobas BCF**, upper trophic, incl. biotransformation	1.0	no	The log Pow of 12.46 is > 9. (The log Pow of 12.46 was estimated by KOWWIN v1.68. The substance is not within the AD of the model.)	BASF SE (2014e)
BCFBAF v3.01 (EPI Suite v4.11): **Arnot-Gobas BCF**, upper trophic, incl. biotransformation of zero	1.2	no	The log Pow of 12.46 is > 9. (The log Pow of 12.46 was estimated by KOWWIN v1.68. The substance is not within the AD of the model.)	BASF SE (2014e)
BCF baseline model v.02.07 (OASIS Catalogic v5.11.13): incl. **mitigating** factors	7.4	no	The substance is within the parametric and the mechanistic, but not within the structural domain due to unknown fragments.	BASF SE (2014f)
CAESAR v2.1.13 (VEGA v1.0.8)	8.0	no	No similar compounds in the training set; accuracy of prediction for similar molecules not optimal; some atom centered fragments not in training set or rare; descriptors with values outside range of training set.	BASF SE (2014b)
BCF baseline model	12.0	no	The substance is within the	BASF SE

v.02.07 (OASIS Catalogic v5.11.13): not considering mitigating factors			parametric and the mechanistic, but not within the structural domain due to unknown fragments.	(2014f)
BCFBAF v3.01 (EPI Suite v4.11): **Meylan** et al. (1997/1999)	28.2	no	The log Pow of 12.46 exceeds upper limit of training set. (The log Pow of 12.46 was estimated by KOWWIN v1.68. The substance is not within the AD of the model.)	BASF SE (2014e)
BCF Read-Across v1.0.2 (VEGA v1.0.8)	44.0	no	Low similarity in found molecules	BASF SE (2014d)
US EPA T.E.S.T. v4.1: Bioaccumulation: Consensus method	101.9	yes, but confidence is low	Results only available from 3 out of 5 models; based on the mean average error, the confidence in the predicted values is low.	BASF SE (2014g)
• Hierarchical clustering	1666.2	yes, but confidence is low		
• FDA	9.7	yes, but confidence is low		
• Nearest neighbor	65.3	yes, but confidence is low		
Meylan v1.0.2 (VEGA v1.0.8)	119.0	no	Only moderately similar compounds with known experimental value in the training set; similar molecules have experimental values that strongly disagree with the target compound predicted value; reliability of log Pow value used by the model is not adequate.	BASF SE (2014c)
Müller and Nendza (2011): Comparative analysis (UBA)			According to the report, the models give inaccurate estimates for compounds with log Pow > 5.	BASF SE (2014a)
Bintein et al. (1993)	< 1	no	log Pow out of range	
European Communities (2003)	< 1	no	log Pow out of range	
Könemann and van Leeuwen (1980)	< 1	no	log Pow out of range; substance not a chlorobenzene; very small training data set	
Connell and Hawker (1988)	179	no	log Pow out of range	
Nendza (1991)	4.51E+04	no	log Pow out of range	
Neely et al. (1974)	9.51E+06	no	log Pow out of range; substance no halogenated aromatics; very small training data set	
[29] Zok et al. (1991)	2.13E+08	no	log Pow out of range; substance not a substituted aniline; very small training data set	
Schüürmann and Klein (1988)	1.60E+09	no	log Pow out of range; substance not a chlorinated or polycyclic hydrocarbon	
Veith and Kosian (1983)	4.30E+09	no	log Pow out of range; substance not a halogenated compound	
Veith et al. (1979)	1.24E+10	no	log Pow out of range	
Escuder-Gilabert et al.	1.58E+10	no	log Pow out of range	

(2001)				
Lu et al. (1999)	4.27E+10	no	log Pow out of range	
Mackay (1982)	2.40E+11	no	log Pow out of range; substance not a chlorinated hydrocarbon	

One experimental study with the substance is available. The guideline study determined a maximum BCF of 1.5 (Kyushu Chemical Biotesting Center, 1986). However, this study must be regarded as invalid as the test concentrations were prepared far above the limit of water solubility; therefore a reliable BCF could not be measured. As the solvent significantly altered the dissolved concentrations in the medium, the study is not valid compared with the recent OECD 305 guideline (2012). Details are given in Chapter 5.3.1.

Therefore, the bioaccumulation potential of 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) has been assessed in a weight of evidence approach due to the lack of valid bioaccumulation testing data. The bioaccumulation was assessed using various scientifically validated QSAR models. However, as the substance is characterised by a complex structure and a very high log Kow, the substance did not comply with the demands of the available models. Nevertheless, depending on the degree of the criteria violations, the estimated BCF values can be used in the assessment of the bioaccumulation potential in combination with other data in a weight-of-evidence approach, e. g. log Pow and water solubility.

In addition to the estimated BCF values, data from a toxicokinetic study have been consulted to assess the potential oral or dermal absorption of mammals regarding the test substance (CTL, 2002) (for details see Chapter 4.1). Following a 6-hour topical exposure, the dermal absorption of 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) was found to be very low and accounted for not more than 0.8% and 0.4% of the applied dose. Furthermore, the topically applied 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) did not achieve systemically measurable concentrations and was thus not bioavailable.
These results are as expected considering the physico-chemical properties of 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol, i.e. the very low water solubility of the test substance (< 5 ng/L) and the high log Pow (>>4). In line, systemic availability of 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) was negligible after oral administration. The test substance 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) was quantitatively and rapidly excreted as parent compound via the faeces.
Based on this information, it could be demonstrated that the substance is not bioavailable as it does not significantly cross biological membranes. Therefore, a significant bioaccumulation in fish is not expected either. This assumption is supported by the QSAR calculations which have been performed using several models.

Table 14 lists the models, the estimated BCF values and basic information on the applicability domain (AD). Detailed information on the model's requirements and the methods are compiled in the (Q)SAR Model Reporting Format (QMRF) of the OECD (Annex 1). Information on the prediction and the criteria of the AD are given in Annex 2.

The estimated BCF values range from less than 1 to 2.40E+11, while the extremely high BCF values were calculated by simple models which do not consider other substance's properties, e.g. ionization or adapt regression equations depending on the range of the log Kow. The substance does not fulfil the requirements of the applicability domain of all models, except for US EPA T.E.S.T. v.4.1 (with low confidence).

5.3.1 Measured bioaccumulation data

In a guideline study investigating the bioaccumulation of the substance in *Cyprinus carpio*, a BCF of maximum 1.5 was determined (Kyushu Chemical Biotesting Center, 1986). However, this study must be regarded as invalid as the test concentrations were prepared far above the limit of water solubility (< 5 ng/L at 20 °C) by a factor of 20,000 (0.1 mg/L) and 200,000 (1 mg/L); therefore a reliable BCF could not be measured. The high test concentrations were selected based on the requirements of the Ministry of Trade and Industry, Japan (MITI). According to the recent OECD guideline 305 (2012) the use of solvent is only accepted at concentrations which do not significantly alter the maximum dissolved concentration in the medium. Regarding the factors between test concentrations and limit of water solubility, this was not the case in the present study. Therefore, the study cannot be regarded as valid in order to determine the BCF in fish.

5.3.2 Estimated bioaccumulation data

5.3.2.1 EPI Suite v4.11: BCFBAF v3.01

Check for OECD Principles for (Q)SAR validation

Defined endpoint	Yes (see Annex 1 for details)
Unambiguous algorithm	Yes (see Annex 1 for details)
Defined domain of applicability	Yes (see Annex 1 for details)
Appropriate measures of goodness-of-fit, robustness and predictivity	Yes (see Annex 1 for details)
Mechanistic interpretation, if possible	Not applicable

The BCFBAF v3.01 program of EPI Suite v4.11 estimates the BCF according to two methods: Meylan et al. (1997/1999) and Arnot-Gobas (2003). For details on the methods see Annex 1, Chapter 1.1. For details on the fulfilment of criteria of the applicability domain see Annex 2, Chapter 9.1.
The Meylan method calculates the BCF based on the log Kow. For non-ionic compounds, one of three algorithms are used to estimate the BCF depending on the log Kow. The regression methodology includes derivation of correction factors based on specific structural features. Regarding CAS 103597-45-1, the BCF was estimated at 28 indicating that significant accumulation in organisms is not to be expected.
However, the maximum log Kow of the training and validation data sets of 11.26 was exceeded; therefore, the substance does not fulfil the requirements of the applicability domain of the model[a]. Nevertheless, as this limit value is relatively close to the substance's log Kow, the estimated BCF can be used in context with other information.
The Arnot and Gobas method restricts the estimation of BCFs to substances with a log Kow of ≤ 9; otherwise the estimate may be highly uncertain. The model calculates a BCF of 1.0 for the upper trophic level considering biotransformation and a BCF of 1.2 without considering biotransformation. These values also indicate that significant accumulation in organisms is not to be expected.

[a] Currently there is no universally accepted definition of model domain. However, users of the model may wish to consider the possibility that bioconcentration factor estimates are less accurate for compounds outside the MW and log Pow ranges of the training set compounds

5.3.2.2 VEGA v1.0.8: CAESAR v2.1.13, Read-Across v1.0.2, Meylan v1.0.2

Check for OECD Principles for (Q)SAR validation: CAESAR v2.1.13, Read-Across v1.0.2, Meylan v1.0.2

Defined endpoint	Yes (see Annex 1 for details)
Unambiguous algorithm	Yes (see Annex 1 for details)
Defined domain of applicability	Yes (see Annex 1 for details)
Appropriate measures of goodness-of-fit, robustness and predictivity	Yes (see Annex 1 for details)
Mechanistic interpretation, if possible	Not applicable

The VEGA platform v1.0.8 combines three models: CAESAR v2.1.13, Read-Across v1.0.2, and Meylan v1.0.2. Details on the method of CAESAR are described in Chapter 1.2.1 (Annex 1), the fulfilment of the applicability domain criteria can be viewed in Chapter 9.2.1 (Annex 2). The substance is not within the applicability domain of CAESAR as no similar compounds were found in the training set. Therefore, the accuracy of prediction was too low. The predicted BCF was 8.
According to the Read-Across model the BCF is 44. However, the similarity of the molecules was low; therefore the substance was not in the applicability domain of the model. Details on the Read-Across method are described in Chapter 1.2.2 (Annex 1), the fulfilment of the applicability domain criteria can be viewed in Chapter 9.2.2 (Annex 2).
The Meylan model predicts a BCF of 119. Again the similarity of compounds in the training set is only moderate. In addition experimental values of these compounds strongly disagree with the predicted BCF. Therefore, the substance is not within the applicability domain of the model. Details on the method of Meylan are described in Chapter 1.2.3 (Annex 1), the fulfilment of the applicability domain criteria can be viewed in Chapter 9.2.3 (Annex 2).
All estimated BCF values indicate that significant accumulation is not to be expected.

5.3.2.3 US EPA T.E.S.T. v4.1: Bioaccumulation factor

Check for OECD Principles for (Q)SAR validation

Defined endpoint	Yes (see Annex 1 for details)
Unambiguous algorithm	Yes (see Annex 1 for details)
Defined domain of applicability	Yes (see Annex 1 for details)
Appropriate measures of goodness-of-fit, robustness and predictivity	Yes (see Annex 1 for details)
Mechanistic interpretation, if possible	Not applicable

The US EPA T.E.S.T. v4.1 model calculates the BCF with the Consensus method which uses the reasonable results of up to five BCF models which estimate BCF values according to a variety of molecular descriptors. T.E.S.T. checks if the substance falls within the applicability domain (AD) of each BCF model and only displays the results of those models if the criteria for the AD are fulfilled. Details on the methods are described in Chapter 1.3 (Annex 1). In case of the substance at hand, only three models produced a BCF within the applicability domain:

- Hierarchical clustering: BCF = 1666.2
- FDA: BCF = 9.7
- Nearest neighbour: BCF = 65.3
- The Consensus method combines these values to a BCF of 101.9.

Although the substance complied with the AD restrictions of the models, the confidence in the estimated BCF is low based on the comparison of the mean absolute error for the complete dataset

with a restricted dataset which only contains substances with a similarity coefficient of 0.5 or higher. Details on the applicability domain and the confidence level can be viewed in Chapter 9.3 (Annex 2).
The calculated BCF of the Consensus method indicates that significant accumulation in organisms is not to be expected.

5.3.2.4 CATALOGIC v5.11.13: BCF base-line model v02.07

Check for OECD Principles for (Q)SAR validation

Defined endpoint	Yes (see Annex 1 for details)
Unambiguous algorithm	Yes (see Annex 1 for details)
Defined domain of applicability	Yes (see Annex 1 for details)
Appropriate measures of goodness-of-fit, robustness and predictivity	Yes (see Annex 1 for details)
Mechanistic interpretation, if possible	Not applicable

The BCF base-line model (v02.07) of OASIS Catalogic (v5.11.13) calculates the BCF based on the substance's structure and its log Kow. It also considers potential mitigating factors such as water solubility, molecular size and metabolism. 2,2'-Methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) (CAS 103597-45-1) completely fulfils the criteria for the parametric and the mechanistic applicability domain. However, the substance contains structural fragments that are unknown to the model (35 % known; 65 % unknown; Table 15). Therefore, the substance is not completely within the applicability domain of the BCF base-line model.
The maximum BCF[b] is estimated at 12.0. The bioaccumulation potential is reduced to a BCF of 7.4 mainly through metabolism and water solubility. The poor water solubility has the highest mitigating effect on the bioaccumulation potential of CAS 103597-45-1 (Table 15). Although the substance is a relatively large molecule as seen by the values for the maximum diameter (DiamMax; see Table 15), its effect on the bioaccumulation potential is rather low, although the PBT Working Group discussed a cut-off value of 17.4 Å for bioaccumulative substances.
Both BCF values – the BCF_{max} and the corrected BCF including mitigating factors - indicate that significant accumulation in organisms is not to be expected.

[b] BCF without considering mitigating factors

Table 15: BCF-baseline v02.07: Model output for CAS 103597-45-1

Model domain similarity	
Parametric domain	In domain
Structural domain	35 % correct
	0 % incorrect
	65 % unknown
Mechanistic domain	In domain
Effects of mitigating factors on BCF	
Acids	0.0000
Metabolism	0.0087
Phenols	0.0000
Size	0.0001
Water solubility	0.0904
Molecular dimensions	
DiamMax-Min [Å]	18.1
DiamMax-Max [Å]	22.8
DiamMax-Mean [Å]	20.2
Estimation	
Log BCF	0.8710±0.1110
BCF	7.4

5.3.2.5 Comparative analysis of estimated and measured BCF data (UBA models: Müller & Nendza, 2011)

Check for OECD Principles for(Q)SAR validation

Defined endpoint	Yes (see Annex 1 for details)
Unambiguous algorithm	Yes (see Annex 1 for details)
Defined domain of applicability	Yes (see Annex 1 for details)
Appropriate measures of goodness-of-fit, robustness and predictivity	Yes (see Annex 1 for details)
Mechanistic interpretation, if possible	Not applicable

Müller and Nendza (2011) compiled 15 regression-based models which rely on the log Kow of which 13 are based on fish bioaccumulation data. Due to the substance's high log Kow, CAS 103597-45-1 does not meet the limits set by the log Kow range of the training sets of the models. In addition, some of the models were based on other substance classes (e. g. chlorobenzenes) and are therefore not suited to estimate a BCF for the substance in question due to a low similarity between the substance and the training set. Some of the models were developed on a very small database (n < 10) and should therefore be regarded as not reliable. The results show a wide BCF range from less than 1 to 2.40E+11 suggesting a low reliability as no trend of the bioaccumulation potential can be derived. This is supported by the report of Müller & Nendza (2011), which found out that the models give inaccurate estimates for a variety of compounds with a log Kow > 5. Details on the methods are described in Chapter 1.5 (Annex 1), the fulfilment of the applicability domain criteria can be viewed in Chapter 9.5 (Annex 2).

5.3.3 Summary and discussion of aquatic bioaccumulation

Due to the lack of experimental data the bioaccumulation potential has been assessed in a weight of evidence approach.

Based on the very low water solubility (< 5 ng/L) experimental BCF studies are technically not feasible. Furthermore, the substance does not fulfil the requirements of the applicability domain of the applied QSAR-models and therefore are not valid, which is mainly due to the substance's structure and its high log Kow (12.7).

Nevertheless, a toxicokinetic study demonstrated that the substance is not bioavailable as it does not significantly cross biological membranes.

In conclusion, the low bioavailability, the poor water solubility, and the high log Kow indicate, that bioaccumulation of the test item in organisms is not to be expected.

5.4 Aquatic toxicity

Table 16: Summary of relevant information on aquatic toxicity

Method	Results	Reliability	Reference
Short-term toxicity to fish – Official Journal of the European Communities L251 (comparable to OECD 203)	LC_{50} (96h) > 28.9 mg/L (measured)	1	CIBA-GEIGY Ltd. (1991b)
Short-term toxicity to aquatic invertebrates - Official Journal of the European Communities L251 (comparable to OECD 202)	LC_{50} (48h) > 65.9 mg/L (measured)	1	CIBA-GEIGY Ltd. (1991a)
Long-term toxicity to aquatic invertebrates (OECD 211)	NOEC (21d) ≥ 25 µg/L (measured)	1	RCC Ltd. (2006)
Long-term toxicity to fish	Not available		
Toxicity to aquatic algae (OECD 201)	EC_{50} (72h) > 2 mg/L (measured) NOEC (72h) ≥ 2 mg/L (measured)	1	Safepharm Laboratories Limited (1995)

5.4.1 Fish

5.4.1.1 Short-term toxicity to fish

A static 96 h freshwater toxicity test was conducted according to the Official Journal of the European Communities L251,vol.27,C-01, 19-09-1984 (comparable to OECD 203) to determine the acute toxicity of the test item to zebra-fish (*Danio rerio*, reported as: *Brachydanio rerio*) (Ciba Geigy Ltd., 1991b). 0.4 % lecithine was used as emulsifier.

At test termination, a LC_{50} >28.9 mg/L (measured) was determined (analytic method: HPLC and UV detector) which complies with the highest measured test concentration under exposure conditions and is clearly above the water solubility of 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol). Thus, no effects in the range of the water solubility could be detected and the test substance can therefore be considered as not harmful to fish.

5.4.1.2 Long-term toxicity to fish

No data available

5.4.2 Aquatic invertebrates

5.4.2.1 Short-term toxicity to aquatic invertebrates

A static 48 h freshwater toxicity test was conducted to determine the acute toxicity of the test item to the water flea *Daphnia magna* according to the Official Journal of the European Communities L251,vol.27,C-01, 19-09-1984 (comparable to OECD 202) (CTL, 2002). 0.4 % lecithine was used as emulsifier.

At test termination, an EC_{50} > 65.9 mg/L (measured) was determined (analytic method: HPLC and UV detector) which complies with the highest attainable concentration and is clearly above the water solubility of 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol). Thus, no effects in the range of the water solubility could be detected and the test substance can therefore be considered as not harmful to aquatic invertebrates.

5.4.2.2 Long-term toxicity to aquatic invertebrates

The effects of the substance on the survival and reproduction of *Daphnia magna* were investigated in a 21 d test which was conducted according to OECD guideline 211 (RCC Ltd, 2006).

In this semi-static test, the test media were renewed three times per week. Due to the low water solubility of the test item, the test media were prepared before the start of the test and prior to each test medium renewal. No auxiliary solvent or emulsifier was used. The measured concentrations of the test item in the freshly prepared test media of the highest test concentration (undiluted filtrate) ranged from < LOQ (limit of quantification of 0.2 µg/L) to 73 µg/L. At the end of the renewal periods, concentrations of the test item between 2.4 and 33 µg/L were measured. There was no significant difference between the concentration measured in samples taken from the actual test at the end of the renewal periods and the concentration measured in samples which were incubated under the test conditions without food and daphnids in parallel to the test. The time-weighed mean concentration (calculated using the concentrations measured at the start and the end of two renewal intervals of 48 hours and one renewal interval of 72 hours) was 25 µg/L at the highest test concentration (undiluted filtrate). The biological results were based on the time-weighted mean concentration of the test item. Taking into account the survival and reproduction of the test animals, which were not affected by the test item up to and including the highest test concentration (undiluted filtrate), the highest concentration of the test item tested without toxic effects after the exposure period of 21 days (21-day NOEC) was at least 25 µg/L. Higher concentrations of the test item could not be tested due to the low water solubility of the test item.

In conclusion, the test item had no toxic effects on survival and reproduction of the daphnids up to the solubility limit of the test item in the test water.

5.4.3 Algae and aquatic plants

The effect of the test item on the growth of the algal species *Scenedesmus subspicatus* over a 72 hour static exposure period was assessed according to OECD guideline 201 (Safepharm Laboratories Ltd, 1995). 0.2 mL/L Tween 80 – tetrahydrofuran was used as emulsifier.

After 72 h an EC_{50} (growth rate) > 2 mg/L (measured) was determined (analytical method: HPLC) which complies with the highest measured test concentration under exposure conditions and is clearly above the water solubility of 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol). The corresponding NOEC is ≥ 2 mg/L (measured). Thus, no effects in the range of the water solubility could be detected and the test substance can therefore be considered as not harmful to algae.

5.4.4 Other aquatic organisms (including sediment)

No data available.

5.5 Comparison with criteria for environmental hazards (sections 5.1 – 5.4)

Environmental hazard criteria according to Regulation (EC) No 1272/2008 – Environmental category Chronic 4 is applied in case when acute or chronic toxicity data do not allow classification but there is still some reason for concern. This category shall be applied in case of:

- poorly water soluble substances (normally < 1 mg/L) which do not reveal acute toxicity at levels up to the water solubility
  AND
- if a substance has the potential to bioaccumulate (BCF ≥ 500 or, if absent, logPow ≥ 4)
  AND
- is also not rapidly degradable.

Comparison of 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) with criteria for environmental hazards – 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) is not rapidly degradable. 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) has a very low water solubility (< 5 ng/L) and shows no acute aquatic toxicity in fish, aquatic invertebrates and algae. The substance is also not toxic in long-term for aquatic invertebrates or algae up to its water solubility limit. For fish there is no long-term toxicity test available. Neither an experimental nor a calculated BCF could be determined. Based on the very low water solubility (< 5 ng/L) and extremely high logPow (12.7) the bioaccumulation potential is expected to be very low. According to ECHA Guidance R.11 "indicators for low uptake could include the lack of observed skin permeability, a very low uptake in long-term mammalian studies, and/or low chronic systemic toxicity in long term mammalian and/or ecotoxicity studies." The oral and dermal toxicokinetic data shows low dermal and oral absorption in rats. This combined with the very low water solubility and the extremely high logPow indicate that there is a very low potential to bioaccumulate. Therefore, 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) does not fulfil the criteria for the environmental hazard category chronic 4.

5.6 Conclusions on classification and labelling for environmental hazards (sections 5.1 – 5.4)

Conclusion of environmental classification according to Regulation (EC) No 1272/2008

According to Part IV of Regulation (EC) No 1272/2008, a substance does not meet the criteria for classification Chronic 4 in case it has no acute or chronic toxicity to algae, aquatic invertebrates or fish up to the limit of water solubility and the substance is not bioaccumulative. Therefore, 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol) should no longer be classified as Aquatic Chronic 4 according to the environmental hazard classification criteria of Regulation (EC) No 1272/2008.

6 OTHER INFORMATION

None

7 REFERENCES

Data have been taken from BASF SE; IUCLID 5; 2,2'-methylenebis(6-(2H-benzotriazol-2-yl)-4-(1,1,3,3-tetramethylbutyl)phenol); 27.10.2011:

BASF SE (2014a). BCF calculated according to 13 models evaluated by Müller & Nendza (2011): CAS no. 103597-45-1. Unpublished. ECT Oekotoxikologie GmbH, Flörsheim, Germany; 23 Sep 2014.

BASF SE (2014b). Calculation of BCF with CAESAR (v2.1.13): CAS 103597-45-1. Unpublished. ECT Oekotoxikologie GmbH, Flörsheim, Germany; 22 Sep 2014.

BASF SE (2014c). Calculation of BCF with Meylan (v1.0.2): CAS 103597-45-1. Unpublished. ECT Oekotoxikologie GmbH, Flörsheim, Germany; 22 Sep 2014.

BASF SE (2014d). Calculation of BCF with Read-Across (v1.0.2): CAS 103597-45-1. Unpublished. ECT Oekotoxikologie GmbH, Flörsheim, Germany; 22 Sep 2014.

BASF SE (2014e). EPI Suite calculation for CAS 103597-45-1. Unpublished. ECT Oekotoxikologie GmbH, Flörsheim, Germany; 19 Sep 2014.

BASF SE (2014f). OASIS Catalogic v5.11.13, BCF base-line model v02.07: BCF predictions for CAS 103597-45-1. Unpublished. BASF SE, Dept. for Product Safety, Ludwigshafen/Rhein, Germany; 19 Sep 2014.

BASF SE (2014g). US EPA T.E.S.T. (v4.1): BCF predictions for CAS 103597-45-1. Unpublished. ECT Oekotoxikologie GmbH, Flörsheim, Germany; 23 Sep 2014.

Bintein S, Devillers J, Karcher W. 1993. Nonlinear Dependence of Fish Bioconcentration on n-Octanol/Water Partition Coefficient. SAR QSAR Environ. Res. 1: 29-39.

Central Toxicology Laboratory (CTL) (2002). Metabolic fate following oral administration or in vivo dermal application in rats. Testing laboratory: Central Toxicology Laboratory (CTL). Report no.: UR0687. Owner company: BASF SE. Report date: 2002-06-07.

CIBA-GEIGY Ltd. (1991a). Report on the Acute Toxicity Test on Daphnia (Daphnia magna Straus 1820). Unpublished data. Testing laboratory: CIBA-GEIGY Ltd. Owner company: BASF SE. Study number: 918111. Report date: 1991-08-13

CIBA-GEIGY Ltd. (1991b). Report on the Acute Toxicity to Zebra-Fish (Brachydanio rerio). Unpublished data. Testing laboratory: CIBA-GEIGY Ltd. Owner company: BASF SE. Study number: 918110. Report date: 1991-08-13

CIBA-GEIGY Ltd. (1991c). Report on the test for ready biodegradability in the Modified Sturm Test. Unpublished data. Testing laboratory: CIBA-GEIGY Ltd. Owner company: BASF SE. Study number: 918113. Report date: 1991-08-26

Connell DW, Hawker DW. 1988. Use of Polynomial Expressions to describe the Bioconcentration of Hydrophobic Chemicals by Fish. Ecotox. Environ. Saf. 16: 242-257.

Escuder-Gilabert L, Martin-Biosca Y, Sagrado S, Villanueva-Camanas RM, Medina-Hernandez MJ. 2001. Biopartitioning Micellar Chromatography to Predict Ecotoxicity. Analytica Chimica Acta 448: 173-185.

European Communities. 2003. Technical guidance document on risk assessment in support of Commission Directive 93/67/EEC on risk assessment for new notified substances,

Commission Regulation (EC) No 1488/94 on risk assessment for existing substances, Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market. Joint Research Centre, Ispra, Italy: European Commission.

Köneman H, van Leeuwen K. 1980. Toxicokinetics in Fish: Accumulation and Elimination of Six Chlorobenzenes by Guppies. Chemosphere 9: 3-19.

Kyushu Chemical Biotesting Center (1986). Bioaccumulation study in Carp. Unpublished data. Testing laboratory: Kyushu Chemical Biotesting Center. Owner company: BASF SE. Study number: 36079. Report date: 1986-01-08

Lu XX, Tao S, Cao J, Dawson RW. 1999. Prediction of Fish Bioconcentration Factors of Nonpolar Organic Pollutants based on Connectivity Indices. Chemosphere 39: 987-999.

Mackay D. 1982. Correlation of Bioconcentration Factors. Environ. Sci. Technol. 16: 274-278.

Müller M, Nendza M (2011). Comparative analysis of estimated and measured BCF data (OECD 305). Umweltbundesamt-Texte 15, 55 pp.

Neely WB, Branson DR, Blau GE. 1974. Partition Coefficients to Measure Bioconcentration Potential of Organic Chemicals in Fish. Env. Sci. Technol. 8: 1113-1115.

Nendza M. 1991. QSARs of bioconcentration: validity assessment of log Pow/log BCF correlations. In Bioaccumulation in aquatic systems, ed. Nagel, R. and Loskill, R. 43-66. Weinheim: VCH.

RCC Ltd. (2005). Inherent biodegradability in a manometric respirometry test. Unpublished data. Testing laboratory: RCC Ltd. Owner company: BASF SE. Study number: 857658. Report date: 2005-04-21

RCC Ltd. (2006). Effect on survival and reproduction of Daphnia magna in a semi-static test over three weeks. Unpublished data. Testing laboratory: RCC Ltd. Owner company: BASF SE. Study number: A72415. Report date: 2006-12-14

Safepharm Laboratories Limited (1995). Algal inhibition test SPL. Unpublished data. Testing laboratory: Safepharm Laboratories Limited. Owner company: BASF SE. Study number: 787/001. Report date: 1995-05-15

Schüürmann G, Klein W. 1988. Advances in Bioconcentration Prediction. Chemosphere 17: 1551-1574.

Veith GD, Defoe DL, and Bergstedt BV. 1979. Measuring and estimating the bioconcentration factor of chemicals in fish. J.Fish.Board Can. 36: 1040-1048.

Veith GD, Kosian P. 1983. Estimating Bioconcentration Potential from Octanol/Water Partition Coefficients. In: Physical Behaviour of PCBs in the Great Lakes. Mackay D, Paterson S, Eisenreich SJ, Simmons MS (Eds.), Ann Arbor Science Publishers, Ann Arbor, MI, U.S.A.

Zok S, Görge G, Kalsch W, Nagel R. 1991. Bioconcentration, Metabolism, and Toxicity of Substituted Anilines in the Zebrafish (*Brachydanio rerio*). Sci.Tot. Environ. 109/110: 411-421.

8 ANNEX 1: QMRF'S: COMPILATION OF INFORMATION ON APPLIED QSAR MODELS

The information on the models is given according to the (Q)SAR Model Reporting Format (QMRF) following the OECD principles stated in REACH Guidance R.6 (ECHA, 2008).

1.1 QMRF: BCFBAF v3.01 (EPI Suite v4.11)

1.0	QSAR identifier	
1.1	QSAR identifier (title)	BCFBAF for estimation of bioconcentration, bioaccumulation and biotransformation in fish
1.2	Other related models	-
1.3	Software coding the model	BCFBAF v3.01 (EPI Suite v4.11)
2.0	General information	
2.1	Date of QMRF	30 Oct. 2013
2.2	QMRF author and contact details	BASF SE, Department of Product Safety, Ludwigshafen, Germany
2.3	Date of QMRF update(s)	-
2.4	QMRF update(s)	-
2.5	Model developer(s) and contact details	The original BCF estimation methodology used by the original BCFWIN program is described in a document prepared for the U.S. Environmental Protection Agency (Meylan et al., 1997) and published by Meylan et al. (1999). BCFBAF has been expanded to include estimation of the Biotransformation Rate (kM) in fish and estimation of Bioaccumulation Factor (BAF) by the Arnot-Gobas method (Arnot and Gobas, 2003).
2.6	Date of model development and/or publication	1. Bioconcentration factor (BCF): Meylan et al., 1997/1999 2. Biotransformation rate in fish (kM): Arnot et al., 2008a/2008b 3. Arnot & Gobas BAF and steady-state BCF: Arnot and Gobas, 2003
2.7	References to main scientific papers and/or software package	- Arnot JA, Gobas FAPC. 2003. A generic QSAR for assessing the bioaccumulation potential of organic chemicals in aquatic food webs. QSAR and Combinatorial Science 22: 337-345. - Arnot JA, Mackay D, Parkerton TF, Bonnell M. 2008a. A database of fish biotransformation rates for organic chemicals. Environmental Toxicology and Chemistry 27(11), 2263-2270. - Arnot JA, Mackay D, Bonnell M. 2008b. Estimating metabolic biotransformation rates in fish from laboratory data. Environmental Toxicology and Chemistry 27: 341-351. - Meylan, W.M., Howard, P.H, Aronson, D., Printup, H. and S. Gouchie. 1997. "Improved Method for Estimating Bioconcentration Factor (BCF) from Octanol-Water Partition Coefficient", SRC TR-97-006 (2nd Update), July 22, 1997; prepared for: Robert S. Boethling, EPA-OPPT, Washington, DC; Contract No. 68-D5-0012; prepared by: ; Syracuse Research

		Corp., Environmental Science Center, 6225 Running Ridge Road, North Syracuse, NY 13212. - Meylan, WM, Howard, PH, Boethling, RS et al. 1999. Improved Method for Estimating Bioconcentration / Bioaccumulation Factor from Octanol/Water Partition Coefficient. Environ. Toxicol. Chem. 18(4): 664-672 (1999).
2.8	Availability of information about the model	The model is non-proprietary and can be downloaded freely from US EPA.
2.9	Availability of another QMRF for exactly the same model	No (http://qsardb.jrc.it/qmrf/).
3.0	Defining the endpoint	
3.1	Species	The bioconcentration factor, the biotransformation rate as well as the bioaccumulation factor of the uncharged molecule is estimated for fish.
3.2	Endpoint	- Bioconcentration factor (BCF) - Bioaccumulation factor (BAF; at 15 °C) - Biotransformation rate (kM) and half-life
3.3	Comment on the endpoint	Regulation (EC) No 1907/2006 [REACH], Annex 1X, 9.3.2 Bioaccumulation in aquatic species, preferably fish
3.4	Endpoint units	- Bioconcentration factor (BCF): L/kg wet weight - Bioaccumulation factor (BAF): L/kg wet weight - Biotransformation rate (kM): per day (normalised to 10 g fish)
3.5	Dependent variable	- Bioconcentration factor (log BCF) - Bioaccumulation factor (log BAF) - Biotransformation rate (kM) and log bio half-life
3.6	Experimental protocol	The bioconcentration of a substance can be determined according to OECD guideline 305.
3.7	Endpoint data quality	The data used for the model development and improvement was taken from quality-reviewed database (review process described in Arnot & Gobas, 2006).
4.0 to 8.0	See below for information on the individual submodels	
9.0	Miscellaneous information	
9.1	Comments	-
9.2	Bibliography	- Arnot JA, Gobas FAPC. 2003. A generic QSAR for assessing the bioaccumulation potential of organic chemicals in aquatic food webs. QSAR and Combinatorial Science 22: 337-345. - Arnot, JA and Gobas FAPC. 2006. A review of bioconcentration factor (BCF) and bioaccumulation factor (BAF) assessments for organic chemicals in aquatic organisms. Environmental reviews 14(4): 257-297. - Arnot JA, Mackay D, Parkerton TF, Bonnell M. 2008a. A database of fish biotransformation rates for organic chemicals. Environmental Toxicology and Chemistry 27(11), 2263-2270. - Arnot JA, Mackay D, Bonnell M. 2008b. Estimating metabolic biotransformation rates in fish from laboratory data.

		Environmental Toxicology and Chemistry 27: 341-351. - CoHort. 2008. CoStatTM Statistical Software, version 6.311. CoHort Software, 798 Lighthouse Ave. PMB 320, Monterey, CA, 93940, USA (http://www.cohort.com) - US EPA (2012). On-Line BCFBAF Help File.
9.3	Supporting information	-

1. Bioconcentration factor (BCF; Meylan et al., 1997/1999)

4.0	Defining the algorithm	
4.1	Type of model	QSAR
4.2	Explicit algorithm	The compound is classified as either non-ionic or ionic (i.e.; carboxylic acids, sulfonic acids and salts of sulfonic acids, and charged nitrogen compounds (nitrogen with a +5 valence such as quaternary ammonium compounds)). **Non-ionic compounds:** Depending on the log Kow one of three algorithms is used to estimate the BCF. The regression methodology includes derivation of correction factors based on specific structural features. Alg. 1: Log Kow < 1.0: Log BCF = 0.50 Alg.2: Log Kow 1.0 to 7.0: Log BCF = 0.6598 Log Kow - 0.333 + Σ correction factors (n = 396, r^2 = 0.792, Q^2 = 0.78, std dev = 0.511, avg dev = 0.395) Alg. 3: Log Kow > 7.0: Log BCF = -0.49 Log Kow + 7.554 + Σ correction factors (n = 35, r^2 = 0.634, Q^2 = 0.57, std dev = 0.538, avg dev = 0.396) **Ionic compounds:** A BCF is assigned based on the log Kow. - Log Kow < 5.0: log BCF = 0.50 - Log Kow 5.0 to 6.0: log BCF = 1.00 - Log Kow 6.0 to 8.0: log BCF = 1.75 - Log Kow 8.0 to 9.0: log BCF = 1.00 - Log Kow > 9.0: log BCF = 0.50
4.3	Descriptors in the model	- Log Kow - Correction factors for structural features of compound
4.4	Descriptor selection	A dataset of 527 compounds with BCF data was used as the training set for developing the estimation algorithms for bioconcentration and for deriving the correction factors. The BCF Non-Ionic Correction Factors are listed in Appendix E of the On-line Help File.
4.5	Algorithm and descriptor generation	Correction factors: The correction factors were derived for specific structural features.
4.6	Software name and version for descriptor generation	- KOWWIN v1.68 (EPI Suite v4.11): log Kow
4.7	Descriptor/Chemicals ratio	- Descriptors: 1 (ionic); 2 (non-ionic) - Chemicals: 61 (ionic); 466 (non-ionic)
5.0	Defining the applicability domain	
5.1	Description of the applicability domain of the model	- Range of molecular weight of the training set - Range of log Kow of the training set - Structural features

5.2	Method used to assess the applicability domain	-
5.3	Software name and version for applicability domain assessment	-
5.4	Limits of applicability	- Molecular Weights in the Training set (n = 527: 466 non-ionic; 61 ionic compounds = carboxylic acids, sulfonic acids, quats): • Ionic: 68.08 to 991.80 • Non-ionic: 68.08 to 959.17 • Average = 244.0 - Log Kow in the Training set: • Ionic: -6.50 to 11.26 • Non-ionic: -1.37 to 11.26
6.0	Defining goodness-of-fit and robustness	
6.1	Availability of the training set	The complete training and validation data sets can be downloaded from the Internet at: http://esc.syrres.com/interkow/EpiSuiteData.htm Substructure searchable formats of the data can be downloaded at: http://esc.syrres.com/interkow/EpiSuiteData_ISIS_SDF.htm The BCF Non-Ionic and Ionic Compound Training Set is also part of Appendix G of the On-Line Help File.
6.2	Available information for the training set	- CAS number - Chemical name - Chemical class - Type of BCF and test conditions - Molecular weight - SMILES - Log Kow - BCF (experimental, estimated) - Concentration of substance in water (measured, nominal) - Exposure conditions (duration, type, temperature) - Fish information (species, wet weight, lipid content, analysed tissue) - BCF (calculation method) - Reference
6.3	Data for each descriptor variable for the training set	**Log Kow:** BCFBAF estimates a log Kow for every SMILES notation by using the estimation module of the KOWWIN program (which is part of the EPI Suite). BCFBAF also automatically retrieves experimental log Kow values from a database containing more than 13200 organic compounds with reliably measured values. When a SMILES structure matches a database structure (via an exact atom-to-atom connection match), the experimental log Kow value is retrieved and used to predict BCF, BAF and kM rather than the estimated value.
6.4	Data for the dependent variable (response) for the training set	**BCF:** Sources/References for BCF listed in training data set.
6.5	Other information	-

	about the training set	
6.6	Pre-processing of data before modelling	- Single BCF values were selected for each compound (median values were generally selected for compounds with multiple values).
6.7	Statistics for goodness-of-fit	Statistical accuracy for the individual algorithms: Alg. 2: n = 396, r^2 = 0.792, Q^2 = 0.78, std dev = 0.511, avg dev = 0.395 Alg. 3: n = 35, r^2 = 0.634, Q^2 = 0.57, std dev = 0.538, avg dev = 0.396 Statistical accuracy of the training data set (non-ionic plus ionic data): - Correlation coefficient (r^2) = 0.833 - Standard deviation = 0.502 log units - Absolute mean error = 0.382 log units
6.8	Robustness – Statistics obtained by leave-one-outcross-validation	-
6.9	Robustness – Statistics obtained by leave-many-outcross-validation	-
6.10	Robustness – Statistics obtained by Y-scrambling	-
6.11	Robustness – Statistics obtained by bootstrap	-
6.12	Robustness – Statistics obtained by other methods	-
7.0	Defining predictivity	
7.1	Availability of the external validation set	The complete training and validation data sets can be downloaded from the Internet at: http://esc.syrres.com/interkow/EpiSuiteData.htm Substructure searchable formats of the data can be downloaded at: http://esc.syrres.com/interkow/EpiSuiteData_ISIS_SDF.htm Appendix H of the On-Line Help File contains the BCF Estimation Method Validation Dataset.
7.2	Available information for the external validation set	See 6.2
7.3	Data for each descriptor variable for external validation set	See 6.3
7.4	Data for the dependent variable for the external validation set	See 6.4

7.5	Other information about the external validation set	-
7.6	Experimental design of test set	As documented in data set
7.7	Predictivity – Statistics obtained by external validation	Statistical accuracy of the validation data set (n = 158 compounds): - Correlation coefficient (r^2) = 0.82 - Standard deviation = 0.59 log units - Absolute mean error = 0.46 log units
7.8	Predictivity - Assessment of the external validation set	-
8.0	Providing a mechanistic interpretation	
8.1	Mechanistic basis of the model	The model estimates the BCF based on the log Kow as hydrophobicity was found to explain more than 70% of the variation of the bioconcentration potential. The model also accounts for the non-ionic or ionic character of the substances by using different equations. In addition correction factors for certain chemical structures were introduced to improve the accuracy of the BCF predictions.
8.2	A priori or a posteriori mechanistic interpretation	-
8.3	Other information about the mechanistic interpretation	-

2. Biotransformation Rate in Fish (kM)

4.0	Defining the algorithm	
4.1	Type of model	QSAR
4.2	Explicit algorithm	- Multiple-linear regression - Log kM/Half-Life (in days) = 0.30734215*LogKow - 0.0025643319*MolWt - 1.53706847 + Σ(Fi*ni) • LogKow: log octanol-water partition coefficient • MolWt: Molecular Weight • Σ(Fi*ni): summation of the individual Fragment coefficient values (Fi) as listed in Appendix F times the number of times the individual fragment occurs in the structure (ni). • The -1.53706847 is the equation constant. Restrictions of model (Arnot et al., 2008): - The model does not account for any transformation in the gill or the gastrointestinal tract. - The model is also not currently applicable to chemicals that are predominantly ionized at physiological pH. - Urinary excretion and dermal absorption are assumed to be insignificant in comparison to the large volumes of water that are

		exchanged at the surface of the gill. - Considering the nature of the data used in the application of the model, reproductive losses are not included.
4.3	Descriptors in the model	- Log Kow - Correction factors for structural features of compound (Appendix F of On-Line Help File) - Molecular weight
4.4	Descriptor selection	-
4.5	Algorithm and descriptor generation	Algorithm (multiple-linear regression) was performed with CoStat statistical software (CoHort, 2008). Correction factors - Structural fragments based on compounds in training set identified - Fragments with no statistical significance were excluded from the final regression
4.6	Software name and version for descriptor generation	- BCFBAF v3.01 (EPI Suite v4.11): $k_{M,N}$ - KOWWIN v1.68 (EPI Suite v4.11): log Kow
4.7	Descriptor/Chemicals ratio	- Descriptors: 3 - Chemicals: 421
5.0	Defining the applicability domain	
5.1	Description of the applicability domain of the model	- Range of molecular weight of the training set - Range of log Kow of the training set - Structural features
5.2	Method used to assess the applicability domain	-
5.3	Software name and version for applicability domain assessment	-
5.4	Limits of applicability	- Molecular weights in the training set (n = 421): 68.08 to 959.17 (average = 259.75) - Log Kow in the training set (n = 421): 0.31 to 8.70 - The model is also not currently applicable to chemicals that are predominantly ionized at physiological pH (Arnot et al., 2008). - The data set used to develop the model did not include metals or organometals, pigments or dyes, or perfluorinated substances and the model should not be used for these substances.
6.0	Defining goodness-of-fit and robustness	
6.1	Availability of the training set	The complete training and validation data sets can be downloaded from the Internet at: http://esc.syrres.com/interkow/EpiSuiteData.htm Substructure searchable formats of the data can be downloaded at: http://esc.syrres.com/interkow/EpiSuiteData_ISIS_SDF.htm Appendix I of the On-Line Help File contains the kM Biotransformation Estimation Method Training Dataset.
6.2	Available information for the training set	- CAS number - Chemical Name - SMILES - Half-life (log HL; measured and predicted)

		- Log Kow - Molecular weight
6.3	Data for each descriptor variable for the training set	**Log Kow:** BCFBAF estimates a log Kow for every SMILES notation by using the estimation module of the KOWWIN program (which is part of the EPI Suite). BCFBAF also automatically retrieves experimental log Kow values from a database containing more than 13200 organic compounds with reliably measured values. When a SMILES structure matches a database structure (via an exact atom-to-atom connection match), the experimental log Kow value is retrieved and used to predict BCF, BAF and kM rather than the estimated value.
6.4	Data for the dependent variable (response) for the training set	- Arnot kM Database (experimental kM biotransformation rates in fish; Arnot et al., 2008; Appendix I of On-Line Help File of BCFBAF) - Database split into training data set with 421 compounds and validation data set with 211 compounds - Biotransformation half-life (log units, days)
6.5	Other information about the training set	-
6.6	Pre-processing of data before modelling	- No data
6.7	Statistics for goodness-of-fit	Statistical accuracy: - Correlation coefficient (r^2) = 0.821 - Correlation coefficient (Q^2) = 0.753 - Standard deviation = 0.494 log units - Absolute mean error = 0.383 log units
6.8	Robustness – Statistics obtained by leave-one-outcross-validation	-
6.9	Robustness – Statistics obtained by leave-many-outcross-validation	-
6.10	Robustness – Statistics obtained by Y-scrambling	-
6.11	Robustness – Statistics obtained by bootstrap	-
6.12	Robustness – Statistics obtained by other methods	-
7.0	Defining predictivity	
7.1	Availability of the external validation set	The complete training and validation data sets can be downloaded from the Internet at: http://esc.syrres.com/interkow/EpiSuiteData.htm Substructure searchable formats of the data can be downloaded at: http://esc.syrres.com/interkow/EpiSuiteData_ISIS_SDF.htm Appendix J of the On-Line Help File contains the kM

		Biotransformation Estimation Method Validation Dataset.
7.2	Available information for the external validation set	See 6.2
7.3	Data for each descriptor variable for external validation set	See 6.3
7.4	Data for the dependent variable for the external validation set	See 6.4
7.5	Other information about the external validation set	-
7.6	Experimental design of test set	-
7.7	Predictivity – Statistics obtained by external validation	Statistical accuracy (n = 211 compounds): - Correlation coefficient (r^2) = 0.734 - Standard deviation = 0.602 log units - Absolute mean error = 0.446 log units
7.8	Predictivity - Assessment of the external validation set	-
8.0	Providing a mechanistic interpretation	
8.1	Mechanistic basis of the model	Bioaccumulation is the net result of relative rates of chemical inputs to an organism from multimedia exposures (e.g., air, food, and water) and chemical outputs (or elimination) from the organism.
8.2	A priori or a posteriori mechanistic interpretation	-
8.3	Other information about the mechanistic interpretation	-

3. Arnot-Gobas BAF/BCF model

4.0	Defining the algorithm	
4.1	Type of model	QSAR
4.2	Explicit algorithm	The program code for the Arnot-Gobas BAF/BCF model is given in Appendix K of the On-Line Help File of BCFBAF
4.3	Descriptors in the model	- Molecular weight - Chemical structure (SMILES), molecular substructures - Log Kow - Normalized whole-body metabolic biotransformation rate constant ($k_{M,N}$; per day; 10 g fish)
4.4	Descriptor selection	Measured BAF data from Great lakes (northern America) for

		poorly metabolised substances Arnot & Gobas (2006): BCF and BAF
4.5	Algorithm and descriptor generation	- Log Kow (user entered, experimental value from software database or estimated) - Normalised whole-body metabolic biotransformation rate constant ($k_{M,N}$; per day; normalised to 10 g fish)
4.6	Software name and version for descriptor generation	- BCFBAF v3.01 (EPI Suite v4.11): $k_{M,N}$ - KOWWIN v1.68 (EPI Suite v4.11): log Kow
4.7	Descriptor/Chemicals ratio	- Descriptors: 2 - Chemicals: 233 organic chemicals (1398 BCF and 997 BAF values for 176 different fish and aquatic invertebrate species)
5.0	Defining the applicability domain	
5.1	Description of the applicability domain of the model	- For limits of applicability see 5.4
5.2	Method used to assess the applicability domain	-
5.3	Software name and version for applicability domain assessment	-
5.4	Limits of applicability	- Model predictions may be highly uncertain for chemicals that have estimated log KOW values > 9. - The model is not recommended for chemicals that appreciably ionize, for pigments and dyes, or for perfluorinated substances. - BCF and BAF estimated for 10 °C (temperate regions). - The model may not adequately capture biotransformation at the first trophic level for chemicals that are readily biotransformed in invertebrates and plankton.
6.0	Defining goodness-of-fit and robustness	
6.1	Availability of the training set	The complete training and validation data sets can be downloaded from the Internet at: http://esc.syrres.com/interkow/EpiSuiteData.htm Substructure searchable formats of the data can be downloaded at: http://esc.syrres.com/interkow/EpiSuiteData_ISIS_SDF.htm
6.2	Available information for the training set	- Chemical characteristics (CAS #, chemical name, molecular weight and empirical or estimated Kow) - Organism characteristics (species, weight, lipid content, tissue analyzed, gender, age, chemical concentration in organism) - Environmental conditions (water temperature, pH, organic carbon content, water type) - Exposure conditions (exposure duration, total chemical concentration, method of water analysis, exposure route) - Experimental design (flow through, static, renewal, methodology in deriving BCF/BAF) - Primary literature reference
6.3	Data for each descriptor variable for the training set	- The BAF calculations were derived from the parameterization and calibration of the model to a large database of measured BAF values from the Great Lakes (Lake Ontario, Lake Erie and Lake

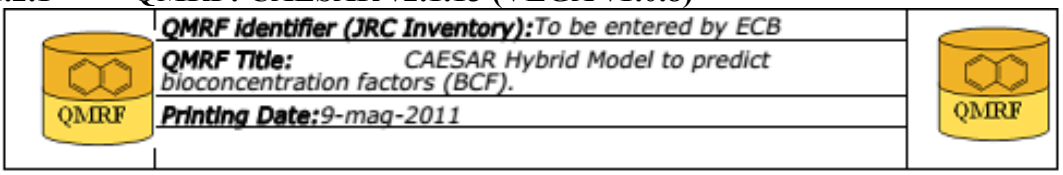
		St. Clair). The measured BAFs are for chemicals that are poorly metabolized (e.g., PCBs) and were generally grouped into lower, middle and upper trophic levels of fish species. - beta: overall food web biomagnification factors in the BAF model are calibrated to each trophic level of measured BAF values - HLN: normalised half-life - The following equations are used to estimate BCF and BAF. For each trophic level BCF and BAF are calculated separately. Tau and the lipid content are the variables which need to be adapted: • Lipid content (Lb): default lipid contents of 10.7%, 6.85% and 5.98% for the upper, middle and lower trophic levels • Bioavailable solute fraction: phi = 1 / (1 + (0.35 * Xpoc * Kow) + (0.08 * Xdoc * Kow)) • Gill uptake rate constant [L kg^{-1} d^{-1}]: $k_1 = 1/((0.01 + 1/Kow) * fish\_wet\_weight^{0.4}$ • Uptake rate constant for chemical in diet [kg kg^{-1} d^{-1}]: $k_D = (0.02 * fish\_wet\_weight^{-0.15} * exp(0.06*T)) / (0.00000005 * Kow + 2)$ • Gill elimination rate constant [d^{-1}]: $k_2 = k_1 / (Lb * Kow)$ • Fecal egestion rate contant [d^{-1}]: $k_E = 0.125 * k_D$ • Growth rate constant [d^{-1}]: $k_G = 0.000502 * pow(fish\_wet\_weight, -0.2)$ • Metabolic biotransformation rate constant [d^{-1}]: $k_M = 0.693/ HLN * pow(fish\_wet\_weight/0.01, -0.25)$ • tau: o upper level: tau = $(0.0065 / (((0.693/HLN) * (0.25/0.01, -0.25)) + 0.0065)^2$ o middle level: tau = $(0.01 / (((0.693/HLN) * (0.03/0.01, -0.25)) + 0.01)^1$ o lower level: tau = $(0.02 / (((0.693/HLN) * (0.016/0.01, -0.25)) + 0.02)^{0.5}$ • ArnotLogBAF = $log10((1-Lb) + (((k_1*phi) + (k_D*beta*phi*tau*Ld*Kow)) /(k_2+k_E+k_G+k_M)))$ • ArnotLogBCF = $log10((1-Lb) + ((k_1*phi) / (k_2+k_E+k_G+k_M)))$ • BAF according to Arnot and Gobas: ArnotBAF = $10^{ArnotLogBAF}$ • BCF according to Arnot and Gobas: ArnotBCF = $10^{ArnotLogBCF}$
6.4	Data for the dependent variable (response) for the training set	See 6.3
6.5	Other information	-

	about the training set	
6.6	Pre-processing of data before modelling	-
6.7	Statistics for goodness-of-fit	No information contained in On-Line Help File. According to Arnot and Gobas (2003), the QSAR produces BAF estimates that are exceeded by only 2.5% of the available empirical data.
6.8	Robustness – Statistics obtained by leave-one-outcross-validation	-
6.9	Robustness – Statistics obtained by leave-many-outcross-validation	-
6.10	Robustness – Statistics obtained by Y-scrambling	-
6.11	Robustness – Statistics obtained by bootstrap	-
6.12	Robustness – Statistics obtained by other methods	-
7.0	Defining predictivity	
7.1	Availability of the external validation set	See 6.1
7.2	Available information for the external validation set	See 6.2
7.3	Data for each descriptor variable for external validation set	See 6.3
7.4	Data for the dependent variable for the external validation set	See 6.4
7.5	Other information about the external validation set	-
7.6	Experimental design of test set	-
7.7	Predictivity – Statistics obtained by external validation	-
7.8	Predictivity -	-

	Assessment of the external validation set	
8.0	Providing a mechanistic interpretation	
8.1	Mechanistic basis of the model	The model includes mechanistic processes for bioconcentration and bioaccumulation such as chemical uptake from the water at the gill surface (BCFs and BAFs) and the diet (BAFs only), and chemical elimination at the gill surface, fecal egestion, growth dilution and metabolic biotransformation (Arnot and Gobas 2003). Other processes included in the calculations are bioavailability in the water column (only the freely dissolved fraction can bioconcentrate) and absorption efficiencies at the gill and in the gastrointestinal tract.
8.2	A priori or a posteriori mechanistic interpretation	-
8.3	Other information about the mechanistic interpretation	-

1.2 VEGA v1.0.8

1.2.1 QMRF: CAESAR v2.1.13 (VEGA v1.0.8)



1.QSAR identifier

1.1.QSAR identifier (title):

CAESAR Hybrid Model to predict bioconcentration factors (BCF).

1.2.Other related models:

Two models, Model A and Model B, have been used to build a hybrid model, Model C. In the proposed approach, the outputs of the individual models (Model A and B) were used as inputs of the final hybrid model.

1.3.Software coding the model:

Freely available in the internet at CAESAR website CAESAR - Computer Assisted Evaluation of industrial chemical Substances According to regulations. coord@caesar-project.eu http://www.caesar-project.eu/software/

2.General information

2.1.Date of QMRF:

21/07/2008

2.2.QMRF author(s) and contact details:

[1]Elena Boriani Istituto di Rcerche Farmacologice Mario Negri boriani@marionegri.it

[2]Manuela Pavan MI&T - Moving Innovation & Technology mpavan@miandt.com

2.3.Date of QMRF update(s):

21/04/2011

2.4.QMRF update(s):

Antonio Cassano
antonio.cassano@marionegri.it
Modified fields: 1.1; 1.2; 1.3; 2.3; 2.4; 2.7; 2.8; 3.3; 3.7; 4.1; 4.2; 4.3; 4.5; 4.6; 5.1; 5.2; 6.2; 6.5; 7.2; 7.5; 7.9; 9.1; 9.2;

Emilio Benfenati
emilio.benfenati@marionegri.it
Modified fields: 1.2; 1.3; 2.2; 2.4; 3.2; 3.3; 4.1; 4.2; 4.4; 4.6; 5.2; 5.4; 9.1; 9.2;

2.5.Model developer(s) and contact details:

[1]Chuyan Zhao Department of Chemistry, Lanzhou University, Lanzhou 730000, China

[2]Elena Boriani Istituto di Ricerche Farmacologiche Mario Negri boriani@marionegri.it http://www.marionegri.it/mn/it/dipLab.html?id=94&ti=4

[3]Antonio Chana Istituto di Ricerche Farmacologiche Mario Negri
[4]Alessandra Roncaglioni Istituto di Ricerche Farmacologiche Mario Negri
aroncaglioni@marionegri.it
http://www.marionegri.it/mn/it/dipLab.html?ti=4&id=549
[5]Emilio Benfenati Istituto di Ricerche Farmacologiche Mario Negri
benfenati@marionegri.it
http://www.marionegri.it/mn/it/dipLab.html?lab=168

2.6.Date of model development and/or publication:

The model was published in 2008.

2.7.Reference(s) to main scientific papers and/or software package:

[1]Zhao, C., Boriani, E., Chana, A., Roncaglioni,A., Benfenati, E. A new hybrid system of QSAR models for predicting bioconcentration factors (BCF). Chemosphere (2008), 73, 1701-1707.
http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6V74-4TSC3V5-1&_user=483112&_coverDate=12%2F31%2F2008&_rdoc=1&_fmt=high&_orig=gateway&_origin=gateway&_sort=d&_docanchor=&view=c&_acct=C000023239&_version=1&_urlVersion=0&_userid=483112&md5=36ee1494fd2e1d3901d6e37e0b368790&searchtype=a
[2]Lombardo A, Roncaglioni A, Boriani E, Milan C, Benfenati E. Assessment and validation of the CAESAR predictive model for bioconcentration factor (BCF) in fish. Chem Cent J. 2010; 4(Suppl 1): S1
http://journal.chemistrycentral.com/content/4/S1/S1

2.8.Availability of information about the model:

A client server application is available to access the model, at http://www.caesar-project.eu

2.9.Availability of another QMRF for exactly the same model:

3.Defining the endpoint - OECD Principle 1

3.1.Species:

Fish (two databases combined; experimental data obtained according to OECD 305 protocol; fish species: Cyprinus Carpio and salmonids).

3.2.Endpoint:

2.Environmental fate parameters 4.Bioconcentration 2.4.a.BCF fish

3.3.Comment on endpoint:

BCF is particularly required under REACH regulation. A good prediction for BCF endpoint may reduce the number of animals (fish) in experimental tests. REACH regulation states that a substance is identified as bioaccumulative (B) when BCF>2000 (log BCF>3.3) and verybioaccumulative (vB) when BCF>5000 (logBCF>3.7). Thus the endpoint could also be treated in classification. Further thresholds apply for the CLP regulation, and for the chemical safety report (CSR), required by REACH. Experimental data are derived by Dimitrov et Al. (see 9.2 bibliography).

3.4.Endpoint units:

BCF unit is l/kg body weight. The modelled variable (logBCF) is adimensional.

3.5.Dependent variable:

LogBCF

3.6.Experimental protocol:

OECD 305 (also standard testing protocol for REACH).

3.7.Endpoint data quality and variability:

Variability of the experimental data: 0.75 log units (Dimitrov et al., 2005), reference in Bibliography, 9.2.

4.Defining the algorithm - OECD Principle 2

4.1.Type of model:

Two models, Model A and Model B, have been used to build hybrid model, Model C. In the proposed approach, the outputs of the individual models (Model A and B) were used as inputs of the hybrid model.

Model A was developed with a Radial Basis Function Neural Network (RBFNN) using an heuristic method to select the optimal descriptors; Model B was developed with a RBFNN using genetic algorithm for the descriptors selection. RBFNN (Wan and Harrington, 1999) was used with a Matlab function for building the models. In-house software made as a PC-Windows Excel macro was used to combine Models A and B within the Model C, using the equations defined in 4.2 (see the supporting information of the Zhao et al. paper in bibliography).

4.2.Explicit algorithm:

The structure of the two RBF NN is implemented in the webtool available at the CAESAR website allowing to reproduce the model. Details about the NN architecture are provided in the supporting information of the paper by Zhao et al. (see 9.2 bibliography). Details of Model A and B are provided in Table1.pdf in 9.3, Supporting information.

If mean (value given by models A and B) > 2.410

log BCF = 1.052 * [mean (value given by models A and B)] - 0.065

If 1.355 < mean (value given by models A and B) ≤ 2.410

log BCF = 0.996 * [min (value given by models A and B)] + 0.042

Otherwise

log BCF = 0.936 * [mean (value given by models A and B)] - 0.123

4.3.Descriptors in the model:

[1]Moriguchi octanol-water partition coefficient (MlogP) Moriguchi et al., 1994

[2]Moran autocorrelation (MATS5V) Molecular descriptor calculated from the molecular graph by summing the products of atom weights of the terminal atoms of all paths of the considered path lenght (the lag)

[3]Number of chlorine atoms (Cl-089) Cl attached to C1(sp2)

[4]Absolute sums of eigenvalues (BEHp2) Molecular descriptor obtained from the positive and negative eigenvalues of the adjacency matrix, weighting the diagonal elements with atom weights.
[5]Geary autocorrelation (GATS5V) Molecular descriptor calculated from the molecular graph by summing the products of atom weights of the terminal atoms of all paths of the considered path lenght (the lag).
[6]X0Solv Solvation connectivity index (X0Solv) Molecular descriptor designed for modelling solvation entropy and describing dispersion interactions in solution.
[7]SsCl Sum of all (-Cl) E-state values in molecule
[8]Aeige Absolute eigenvalues sum from electronegativity weighted distance matrix

4.4.Descriptor selection:

The set of descriptors initially screened is made of 2D molecular descriptors, calculated by DRAGON version 5.4 (759 descriptors), MDL descriptors (249 descriptors), ACD labs version 9.08, (13 descriptors) and KOWWIN (1 descriptor). Thus, 1022 descriptors were obtained including different logP and logD values calculated with these programs. The final, implemented model, available on the web, uses only descriptors calculated with DRAGON 5.4. Heuristic and genetic algorithm methods were used to select the optimal descriptors.

The hybrid model was derived from Model A (HM +RBFNN) and Model B (GA +RBFNN). A heuristic (HM) (Zhao et al., 2008) and genetic algorithm (GA) methods were used to select optimal descriptors. The software CODESSA (see Katrizky et al. 2005 in bibliography 9.2) version 2.21 was used for the HM, to give a complete search for the best multilinear correlations in the ordinary least squares regression (OLS) method. MobyDigs version 1.0 (http://www.talete.mi.it) was used for Genetic Algorithm-Variable Subset Selection strategy (GA-VSS).

4.5.Algorithm and descriptor generation:

2D descriptors have been used. Hybrid model (Model C), combining 2 models: Model A (HM + RBFNN) and Model B (GA + RBFNN).

4.6.Software name and version for descriptor generation:

Codessa 2.21
CODESSA was used to apply HM for variable selection.
support@semichem.com
http://www.semichem.com/codessa/default.php

Moby Digs 1.0
software for selection of variables by Genetic Algorithms
info@talete.mi.it
http://www.talete.mi.it

DRAGON version 5.4

software for calculation of molecular descriptors
info@talete.mi.it
http://www.talete.mi.it

ACD labs 9.08
software for calculation of logP and logD.
http://accelrys.com/products/databases/sourcing/available-chemicals-directory.html

Kowwin 1.67
Estimates the log octanol-water partition coefficient, logP, of chemicals using an atom/fragment contribution method.
howardp@syrres.com
http://www.epa.gov/opptintr/exposure/pubs/episuite.htm

MDL
software for calculation molecular descriptors
http://accelrys.com/products/databases/sourcing/available-chemicals-directory.html

4.7.Descriptors/Chemicals ratio:

378 chemicals in the training set / 8 descriptors = 47.25

5.Defining the applicability domain - OECD Principle 3

5.1.Description of the applicability domain of the model:

The BCF data set is characterised by chemicals of broad nature, with a good presence of hydrocarbons and halogenated compounds, containing many chemicals with single functional groups in a high percentage.

The users have three different ways to evaluate the applicability domain of the model provided by CAESAR:

1) Descriptors Range checked automatically; if a descriptor is out of range, an error message happens:

The model is suitable for compounds that have the descriptors in the following ranges:

MLOGP: min -1,54; max 8,41
X0sol: min 3,54; max 29,34
MATS5v: min -3,93; max 3
BEHp2: min 1,24; max 5,22
SsCl: min 0; max 51,01
AEige: min 4,23; max 534,33
GATS5v: min 0; max 6,93
Cl-089: min 0; max 6

2) CAESAR Remarks: CAESAR software shows a remark if a fragment related to chemical identified as outlier is found (see Lombardo et al. in bibliography 9.2).

3) Similar Compounds:

CAESAR application visualizes the six most similar compounds in the training/test set. The user should check if the similarity is at least > 0.7 for one of the six similar chemicals and the behavior of the model in estimating the similar compounds.

5.2.Method used to assess the applicability domain:

Within CAESAR a special tool was developed. This tool, available at the website (http://www.caesar-project.eu/), shows the six most similar compounds present in our data set, and the related experimental and predicted values. In this way the user can have a direct, transparent, and clear assessment of the errors for similar compounds, and thus have a good basis for the evaluation of the applicability domain specific for a certain compound. Indeed, this information is related to the compound of interest. Moreover, CAESAR shows remarks about the presence of fragments related to chemicals identified as outliers. Finally CAESAR visualizes a warning if the range of calculated descriptors for a single compound is different from those on training set.

5.3.Software name and version for applicability domain assessment:

CAESAR v.1.0

The CAESAR Application is a JAVA™ web application that allows the access to all the toxicity predictive models developed within the CAESAR Project.

coord@caesar-project.eu

http://www.caesar-project.eu/software/

5.4.Limits of applicability:

It is not possible to process with CAESAR model inorganic compounds, mixtures (in addition consider that stereoisomers are not distinguished) and metal complexes. Salts are treated in their neutralized form (free acid).

6.Internal validation - OECD Principle 4

6.1.Availability of the training set:

Yes

6.2.Available information for the training set:

CAS RN:Yes

Chemical Name:Yes

Smiles:Yes

Formula:Yes

INChI:No

MOL file:Yes

6.3.Data for each descriptor variable for the training set:

All

6.4.Data for the dependent variable for the training set:

All

6.5.Other information about the training set:

The whole training set is provided in supporting information (Training_set.xls).

The training set structures are provided in supporting information (structures_training.sdf)

6.6.Pre-processing of data before modelling:

All chemical structures have been double-checked manually.

6.7.Statistics for goodness-of-fit:

Full details on the statistics are in Zhao et al., 2008 (see 9.2 bibliography). Briefly, R^2 (training set) = 0,83. Furthermore, as classifier, the prediction accuracy was 98%, sensitivity 96%, specificity 100%.

6.8.Robustness - Statistics obtained by leave-one-out cross-validation:

6.9.Robustness - Statistics obtained by leave-many-out cross-validation:

Leave many out (20%) cross validation models (20% of the compounds on the training set were randomly selected (sub-test set) and a model developed with the remaining ones (sub-training set). This procedure was repeated 10 times. Results is: Rcv2= 0.79 , SDEP = 0.66

6.10.Robustness - Statistics obtained by Y-scrambling:

6.11.Robustness - Statistics obtained by bootstrap:

6.12.Robustness - Statistics obtained by other methods:

7.External validation - OECD Principle 4

7.1.Availability of the external validation set:

Yes

7.2.Available information for the external validation set:

CAS RN:Yes
Chemical Name:Yes
Smiles:Yes
Formula:Yes
INChI:No
MOL file:Yes

7.3.Data for each descriptor variable for the external validation set:

All

7.4.Data for the dependent variable for the external validation set:

All

7.5.Other information about the external validation set:

The test set is provided in supporting information (Test_set.xls).

The test set structures are provided in supporting information

The splitting of the chemicals has been done keeping into acc
chemical composition, considering the presence of atoms
, nitrogen, etc.

ılictivity - Statistics obtained by external validation:

Full details on the statistics are in Zhao et al., 2008 (see
ıphy). R2 (test set) = 0,80. Only five of the outliers (55, 57,
6) are false negatives (see 5.1 applicability domain).

ılictivity - Assessment of the external validation set:

Further assessment of the model has been done with a se
l set of 527 compounds. Results confirmed the model predict
he external set = 0.81. Full details have been published and
ın be dowloaded at the CAESAR web site (see Lombardo et a
iography).

ıments on the external validation of the model:

The selected substances were split into the training (80% of
ıces) and the test (20% of the substances) sets of the mod

ding a mechanistic interpretation - OECD Principle 5

hanistic basis of the model:

The model largely relies on logP, which typically is the r
:or used for BCF. Corrections are applied to balance the us
cific logP calculator, MLogP. Indeed, this particular descr
ɔod results when chemicals contain C,N,O, but it may be
e in case of compounds with other atoms, like Cl and P.

riori or a posteriori mechanistic interpretation:

The mechanistic interpretation of the model is provid
ɔri, i.e. by interpretation of the final set of the sele
tors.

er information about the mechanistic interpretation:

llaneous information

ıments:

The CAESAR model can be used also in classification (Lombard
2010).
e hybrid model also performed well as a classifier for "B" and
:als. Another important feature of models for regulatory purp
ɔducibility. To obtain that, the parameters of the model hav
d. Within this CAESAR model any user will get exactly the s
when introducing the structure for a given chemical, using
a described before. This shows that the model is reproducible

et al. 2005), all structures were checked one-by-one within the EC funded project CAESAR, by at least two scientists.

9.2.Bibliography:

[1]Dimitrov, S. et al., 2005. SAR QSAR Environ. Res., 16, 531-554
[2]Zhao et al., Chemosphere, Volume 73, Issue 11, December 2008, Pages 1701-1707 http://www.sciencedirect.com/science?_ob=ArticleURL&_udi=B6V74-4TSC3V5-1&_user=483112&_coverDate=12%2F31%2F2008&_rdoc=1&_fmt=high&_orig=gateway&_origin=gateway&_sort=d&_docanchor=&view=c&_acct=C000023239&_version=1&_urlVersion=0&_userid=483112&md5=36ee1494fd2e1d3901d6e37e0b368790&searchtype=a
[3]Wan and Harrington, 1999. J.Chem.Inf.Comput.Sci., 39, 1049-1056.
[4]Katrizky, A.R., et al. (2005). Comprehensive Descriptors for structural and Statistical Analysis. University of Florida. http://www.semichem.com/codessa/ default.php
[5]Moriguchi, L. et al., 1994. Chem Pharm. Bull., 42, 976-978.
[6]Lombardo et al. Chem Cent J. 2010; 4(Suppl 1): S1 http://journal.chemistrycentral.com/content/4/S1/S1

9.3.Supporting information:

Training set(s)Test set(s)Supporting information

10.Summary (ECB Inventory)

10.1.QMRF number:

To be entered by ECB

10.2.Publication date:

To be entered by ECB

10.3.Keywords:

To be entered by ECB

10.4.Comments:

To be entered by ECB

1.2.2 QMRF: BCF Read-Across v1.0.2 (VEGA v1.0.8)

The model performs a read-across on a dataset of 860 chemicals. This dataset has been made by Istituto di Ricerche Farmacologiche Mario Negri (Milan, Italy), merging experimental data from several reliable sources, including the original dataset of the CAESAR BCF model (Zhao et al. 2008, Lombardo et al. 2010; note that experimental values may differ from the ones in the CAESAR BCF dataset, as this new dataset has been built including more sources). The read-across is based on the similarity index developed inside the VEGA platform (accessible athttp://www.vega-qsar.eu/). The index takes into account several structural aspects of the compounds, such as their fingerprint, the number of atoms, of cycles, of heteroatoms, of halogen atoms, and of particular fragments (such as nitro groups). The index value ranges from 1 (maximum similarity) to 0. On the basis of this structural similarity index, the three compounds from the dataset resulting most similar to the chemical to be predicted are taken into account: the estimated BCF value is calculated as the weighted average value of the experimental values of the three selected compounds, using their similarity values as weight.

Estimation Accuracy

Following, statistics obtained applying the read-across prediction to its original dataset, with a leave-one-out approach (read-across for each compound has been performed on the whole dataset without the compound itself)
n = 860; R^2 = 0.63; RMSE = 0.81
Furthermore, the statistics considering the Applicability Domain (AD) index is here reported. The AD index is used to choose only the results that are considered fully reliable predictions (614 over 860 compounds), showing that this subset of compounds has better performance:
n = 614; R^2 = 0.73; RMSE = 0.69

References

- VEGA Guide to BCF Read-Across version 1.0.2 implemented in the VEGA tool v1.0.8
- Zhao, C., Boriani, E., Chana, A., Roncaglioni,A., Benfenati, E. A new hybrid system of QSAR models for predicting bioconcentration factors (BCF).Chemosphere (2008), 73, 1701-1707.
- Lombardo A, Roncaglioni A, Boriani E, Milan C, Benfenati E. Assessment and validation of the CAESAR predictive model for bioconcentration factor (BCF) in fish. Chemistry Central Journal (2010), 4 (Suppl 1).

1.2.3 QMRF: Meylan v1.0.2 (VEGA v1.0.8)

The model is based on the method proposed by Meylan et al. (1999) implemented in the EPI Suite BCFBAF module (http://www.epa.gov/oppt/exposure/pubs/episuite.htm). The model provides a BCF prediction based on different regression equations or fixed values, selected on the basis of an initial classification between ionic and non-ionic compounds, and on the value of the predicted logP value.
For the purpose of the model, ionic compounds include carboxylic acids, sulfonic acids and salts of sulfonic acids, and charged nitrogen compounds (nitrogen with a +5 valence such as quaternary ammonium compounds). All other compounds are classified as non-ionic. The logP prediction is provided by the VEGA logP model.
The original dataset from EPI Suite has been taken, then processed and cleared from duplicates and compounds provided with structure that had problems. The final dataset has 662 compounds.
Non-Ionic compounds
Methodology for Non-Ionic was to separate compounds into three divisions by Log Kow value as follows:
- Log Kow < 1.0
- Log Kow 1.0 to 7.0
- Log Kow > 7.0

For each division, a "best-fit" straight line was derived by common statistical regression methodology. The regression methodology includes derivation of correction factors based on specific structural features. Non-ionic compounds are predicted by the following relationships:
For Log Kow 1.0 to 7.0 the derived QSAR estimation equation is:
Log BCF = 0.6598 Log Kow - 0.333 + Σ correction factors
(n = 396, r^2 = 0.792, Q^2 = 0.78, std dev = 0.511, avg dev = 0.395)
For Log Kow > 7.0 the derived QSAR estimation equation is:
Log BCF = -0.49 Log Kow + 7.554 + Σ correction factors
(n = 35, r^2 = 0.634, Q^2 = 0.57, std dev = 0.538, avg dev = 0.396)
Certain super-hydrophobic chemicals (Log Kow >7.0) selected from the empirical database had reported BCF values with measured water concentrations that exceed water solubility limits. These BCF values were corrected based on estimates of water solubility limits (Arnot and Gobas, 2006).
For Log Kow < 1.0 the derived QSAR estimation equation is: All compounds with a log Kow of less than 1.0 are assigned an estimated log BCF of 0.50.
Ionic compounds

Ionic compounds are predicted as follows:
log BCF = 0.50 (log Kow < 5.0)
log BCF = 1.00 (log Kow 5.0 to 6.0)
log BCF = 1.75 (log Kow 6.0 to 8.0)
log BCF = 1.00 (log Kow 8.0 to 9.0)
log BCF = 0.50 (log Kow > 9.0)
Estimation Accuracy
Following, statistics obtained applying the model to its original dataset:
- Training set: n = 516; R^2 = 0.80; RMSE = 0.55
- Test set: n =146; R^2 = 0.79; RMSE = 0.66
Furthermore, the statistics for the test set considering the Applicability Domain (AD) index is reported here; the AD index is used, as in the final model's assessment, in order to divide results in three groups (into AD, possibly out of AD, out of AD), showing that compounds considered into AD have better performance than the others:
- Test set with AD index greater than 0.85 (compounds into the AD):
- n = 36; R^2 = 0.91; RMSE = 0.45
- Test set with AD index between 0.85 and 0.7 (compounds could be out of AD):
- n = 58; R^2 = 0.79; RMSE = 0.53
- Test set with AD index lower than 0.7 (compounds out of the AD):
- n = 52; R^2 = 0.74; RMSE = 0.87
References

- VEGA Guide to BCF Meylan Model version 1.0.2 implemented in the VEGA tool v1.0.8
- Meylan W.M., Howard PH, Boethling RS et al. 1999. Improved Method for Estimating Bioconcentration / Bioaccumulation Factor from Octanol/Water Partition Coefficient. Environ. Toxicol. Chem. 18(4): 664-672 (1999).
- Arnot J.A. and Gobas F.A.P.C. 2006. A review of bioconcentration factor (BCF) and bioaccumulation factor (BAF) assessments for organic chemicals in aquatic organisms. Environmental reviews 14(4): 257-297.

1.3 QMRF: US EPA T.E.S.T. v4.1: Bioaccumulation factor

1.0	QSAR identifier	
1.1	QSAR identifier (title)	Estimation of bioaccumulation in fish using T.E.S.T. v4.1
1.2	Other related models	-
1.3	Software coding the model	T.E.S.T. v4.1
2.0	General information	
2.1	Date of QMRF	08 July 2014
2.2	QMRF author and contact details	BASF SE, Department of Product Safety, Ludwigshafen, Germany
2.3	Date of QMRF update(s)	-
2.4	QMRF update(s)	-
2.5	Model developer(s) and contact details	US EPA (Todd Martin, Paul Harten, Raghuraman Venkatapathy, and Douglas Young)
2.6	Date of model development and/or publication	2012

2.7	References to main scientific papers and/or software package	User's Guide for T.E.S.T. (version 4.1) (Toxicity Estimation Software Tool). US EPA, 2012.
2.8	Availability of information about the model	The model is non-proprietary and can be downloaded freely from US EPA (http://www.epa.gov/nrmrl/std/qsar/qsar.html)
2.9	Availability of another QMRF for exactly the same model	No (http://qsardb.jrc.it/qmrf/).
3.0	Defining the endpoint	
3.1	Species	The bioconcentration factor is estimated for fish.
3.2	Endpoint	Bioconcentration factor (BCF)
3.3	Comment on the endpoint	Regulation (EC) No 1907/2006 [REACH], Annex 1X, 9.3.2 Bioaccumulation in aquatic species, preferably fish
3.4	Endpoint units	-
3.5	Dependent variable	Bioconcentration factor (log BCF)
3.6	Experimental protocol	The bioconcentration of a substance can be determined according to OECD guideline 305.
3.7	Endpoint data quality	
4.0	Defining the algorithm	
4.1	Type of model	QSAR
4.2	Explicit algorithm	T.E.S.T. uses six methods to estimate the BCF. The results of five methods can be used individually to assess the bioaccumulation potential of a substance, while the sixth method (Consensus) depends upon the output of the other models. Hierarchical clustering The BCF for a given query compound is estimated using the weighted average of the predictions from several different models. The different models are obtained by using Ward's method to divide the training set into a series of structurally similar clusters. A genetic algorithm based technique is used to generate models for each cluster. The hierarchical clustering method produces a series of clusters with similar properties from the training set. in an optimisation procedure, outliers are removed from the clusters and the model building process is repeated. Both processes are repeated until no further outliers are detected. The q^2 LOO (Leave One Out correlation coefficient) must be greater than or equal to 0.5 in order to be valid. The models are generated prior to runtime. The predicted BCF for a test chemical is given by the weighted average for all the valid predictions. In the current version of the software, the predictions are made using the closest cluster from each step in the hierarchical clustering. FDA (Food and Drug Administration) method This method is based on the work of Contrera et al. (2003). Predictions for the chemical in question are made using a unique cluster which contains structurally similar chemicals from the

		overall training set. The unique cluster is constructed at runtime of the model. In this version of the software, clusters are constructed using the thirty most similar chemicals from the training set in terms of the cosine similarity coefficient. A minimum similarity coefficient of 75% is not required. Otherwise no prediction is made. Single model The single model is a single multiple linear regression model using molecular descriptors as independent variables. Techniques and constraints for building the model are similar to those for the hierarchical method with the exception that the single model is fit to the entire training set. The model is generated prior to runtime. The advantage of this method is that a simple transparent model can be developed which does not rely on clustering the chemicals correctly. The disadvantage of this approach is that sometimes an overall model cannot correctly correlate the BCF for every chemical class (Benigni and Richard 1996). Group contribution Method based on group contribution approach of Martin and Young (2001). Fragment counts are used to fit a multiple linear regression model to the entire data set. In order to make a prediction the final model must include at least three molecules in the training set with each fragment of the test chemical, outliers are removed and the process of regression and outlier removal is iterated until no more outliers are found. The regression model is generated prior to runtime. The advantage of this approach is a single transparent model. The disadvantage is that it assumes that the contribution of each fragment does not depend on the presence of nearby fragments in the molecule. Nearest neighbour The predicted BCF is the average of the BCF values of the three most structurally analogues in the training set. The advantage is a quick external estimate of the BCF while the disadvantage is that structural differences between the test chemical and its structural analogues are not accounted for. Consensus This model predicts the BCF by calculating the average of the predicted BCF values from the other QSAR methodologies while taking the applicability domain of the models into account (Zhu et al., 2008). The method is only applied if more than one QSAR model can make a prediction for the substance in question. This method typically provides the highest prediction accuracy since errant predictions are dampened by the predictions from the other methods. In addition this method provides the highest prediction coverage because several methods with slightly different applicability domains are used to make a prediction.
4.3	Descriptors in the model	Molecular descriptors are used to develop the models. The overall pool of descriptors in the software contain 797 2-dimensional descriptors of the following classes: E-state values and E-state counts, constitutional descriptors, topological descriptors, walk and

		path counts, connectivity, information content, 2d autocorrelation, Burden eigenvalue, molecular property (such as the octanol-water partition coefficient), Kappa, hydrogen bond acceptor/donor counts, molecular distance edge, and molecular fragment counts. The descriptors used to describe the compound can be viewed in the model output details.
4.4	Descriptor selection	Not specified
4.5	Algorithm and descriptor generation	The toxicity for a given query compound is estimated using the weighted average of the predictions from several different models. The different models are obtained by using Ward's method to divide the training set into a series of structurally similar clusters. A genetic algorithm based technique is used to generate models for each cluster. The models are generated prior to runtime.
4.6	Software name and version for descriptor generation	The basis of the molecular calculations was the Chemistry Development Kit (Steinbeck et al. 2003). The descriptor values were validated using MDL QSAR (Elsevier MDL 2006), Dragon (Talete 2006), and Molconn-z (Edusoft-LC 2006). The descriptor values were generally in good agreement (aside from small differences in the descriptor definitions for descriptors such as the number of hydrogen bond acceptors).
4.7	Descriptor/Chemicals ratio	The software contains 797 2-dimensional molecular descriptors. The final dataset consists of 676 chemicals.
5.0	Defining the applicability domain	
5.1	Description of the applicability domain of the model	Hierarchical clustering The applicability domain of the cluster models is defined by three constraints: 1) Model ellipsoid constraint: test chemical is within the multidimensional ellipsoid defined by the ranges of descriptor values for the chemicals in the cluster (for the descriptors appearing the cluster model). 2) Rmax constraint: distance from the test chemical to the centroid of the cluster is less than the maximum distance for any chemical in the cluster of the cluster centroid 3) Fragment constraint: the compounds in the cluster have to have at least one example from each of the fragments contained in the test chemical. The fragment constraint can be removed by checking the **Relax fragment constraint** checkbox. FDA (Food and Drug Administration) method The LOO q^2 must be at least 0.5 for a cluster to have a valid predictive model. If the model for the cluster does not satisfy these constraints the cluster size is increased incrementally (maximum size 75 chemicals) until a valid prediction can be made. Otherwise no prediction is made. Single model No specific information is given. Group contribution The constraints for the predictions are similar to the hierarchical method (model ellipse, fragment). Nearest neighbour As a prerequisite the cosine similarity coefficient (SCmin) must be

		greater than or equal to 0.5 Martin et al., 2008). Consensus This method only uses results from valid models (Zhu et al., 2008). No prediction is made if only one valid result is available. The output of the T.E.S.T. only contains results from valid models.
5.2	Method used to assess the applicability domain	-
5.3	Software name and version for applicability domain assessment	-
5.4	Limits of applicability	-
6.0	Defining goodness-of-fit and robustness	
6.1	Availability of the training set	Data was compiled from several different databases (Dimitrov et al. 2005; Arnot and Gobas 2006; EURAS ; Zhao 2008). The final dataset consists of 676 chemicals (after removing salts, mixtures, and ambiguous compounds). http://www.epa.gov/nrmrl/std/qsar/DataSets.zip
6.2	Available information for the training set	Not specified in User's Guide
6.3	Data for each descriptor variable for the training set	Not specified in User's Guide
6.4	Data for the dependent variable (response) for the training set	Not specified in User's Guide
6.5	Other information about the training set	Data provided in sdf format (structure-data file).
6.6	Pre-processing of data before modelling	Salts, mixtures, and ambiguous compounds were removed from the datasets
6.7	Statistics for goodness-of-fit	The predictive ability of each of the QSAR methodologies was evaluated using statistical external validation (Gramatica and Pilutti 2004). Random selection was used to develop the training and test sets. A QSAR model has acceptable predictive power if the following conditions are satisfied (Golbraikh et al. 2003, Journal of Computer-Aided Molecular Design 17, 241 -253.): $q^2 > 0.5$; $R^2 > 0.6$; $(R^2 – R_0^2)/R^2 < 0.1$; $0.85 <= k <= 1.15$ q^2: leave one out correlation coefficient for the training set
6.8	Robustness – Statistics obtained by leave-one-outcross-validation	-

6.9	Robustness – Statistics obtained by leave-many-outcross-validation	-
6.10	Robustness – Statistics obtained by Y-scrambling	-
6.11	Robustness – Statistics obtained by bootstrap	-
6.12	Robustness – Statistics obtained by other methods	-
7.0	Defining predictivity	
7.1	Availability of the external validation set	Random selection was used to develop the training and test sets. See 6.1.
7.2	Available information for the external validation set	-
7.3	Data for each descriptor variable for external validation set	-
7.4	Data for the dependent variable for the external validation set	-
7.5	Other information about the external validation set	-
7.6	Experimental design of test set	-

<table>
<tr><td>7.7</td><td>Predictivity – Statistics obtained by external validation</td><td>
<table>
<tr><th>Method</th><th>R^2</th><th>$(R^2-R_0^2)/R^2$</th><th>k</th><th>RMSE</th><th>MAE</th><th>Coverage</th></tr>
<tr><td>Hierarchical</td><td>0.734</td><td>0.019</td><td>0.888</td><td>0.712</td><td>0.541</td><td>0.926</td></tr>
<tr><td>Single Model</td><td>0.742</td><td>0.083</td><td>0.901</td><td>0.684</td><td>0.543</td><td>0.926</td></tr>
<tr><td>FDA</td><td>0.705</td><td>0.036</td><td>0.905</td><td>0.746</td><td>0.571</td><td>0.911</td></tr>
<tr><td>Group Contribution</td><td>0.675</td><td>0.187</td><td>0.888</td><td>0.760</td><td>0.622</td><td>0.874</td></tr>
<tr><td>Nearest neighbor</td><td>0.609</td><td>0.100</td><td>0.931</td><td>0.884</td><td>0.604</td><td>0.948</td></tr>
<tr><td>Consensus</td><td>0.760</td><td>0.066</td><td>0.900</td><td>0.661</td><td>0.513</td><td>0.926</td></tr>
<tr><td>BCFBF v3.00 (US EPA EPI Suite, 2009)</td><td>0.766</td><td>-</td><td>-</td><td>-</td><td>0.50</td><td>-</td></tr>
</table>

R^2: correlation coefficient between the observed and predicted toxicities for the test set

R_0^2: correlation coefficient between the observed and predicted toxicities for the test set with the y-intercept set to zero (regression line: y=kx)

k: slope of the line y=kx for the test set

RMSE: root mean square error

MAE: mean absolute error

coverage: prediction coverage, fraction of chemicals predicted

In the external statistical evaluation, the consensus method yielded the best results in terms of prediction accuracy and coverage.

For comparison, the statistical values for the widely used BCFBAF v3.00 module of the EPI Suite package are given in the table. The results from the BCFBAF module of EPI Suite are based on the same chemicals that were able to be predicted by the consensus method. The predictions for the consensus method are comparable to those from EPI Suite.</td></tr>
<tr><td>7.8</td><td>Predictivity - Assessment of the external validation set</td><td>-</td></tr>
<tr><td>8.0</td><td colspan="2">Providing a mechanistic interpretation</td></tr>
<tr><td>8.1</td><td>Mechanistic basis of the model</td><td>The mechanistic basis of the models are not provided in detail for every model in the User’s Guide (US EPA, 2012). The BCF is</td></tr>
</table>

		estimated based on molecular descriptors, e.g. fragment counts.
8.2	A priori or a posteriori mechanistic interpretation	-
8.3	Other information about the mechanistic interpretation	-
9.0	Miscellaneous information	
9.1	Comments	-
9.2	Bibliography	- Benigni, R., and Richard, A. M. 1996. QSARS of mutagens and carcinogens: Two case studies illustrating problems in the construction of models for noncongeneric chemicals. Mutation Research 371:29-46. - Contrera, J. F., Matthews, E. J., and Benz, R.D. 2003. Predicting the carcinogenic potential of pharmaceuticals in rodents using molecular structural similarity and E-state indices. Regulatory Toxicology and Pharmacology 38: 243-259. - Gramatica, P., and Pilutti, P. 2004. Evaluation of different statistical approaches for the validation of quantitative structure-activity relationships. Ispra, Italy: The European Commission - Joint Research Centre, Institute for Health & Consumer Protection - ECVAM. - Martin, T. M., Harten, P., Venkatapathy, R., Das, S., and Young, D. M. 2008. A Hierarchical Clustering Methodology for the Estimation of Toxicity. Toxicology Mechanisms and Methods 18:251–266. - Martin, T. M., and Young, D. M. 2001. Prediction of the Acute Toxicity (96-h LC50) of Organic Compounds to the Fathead Minnow (*Pimephales promelas*) Using a Group Contribution Method. Chemical Research in Toxicology 14:1378-1385. - US EPA (2008). Molecular Descriptors Guide – Description of the Molecular Descriptors Appearing in the Toxicity Estimation Software Tool. Version 1.0.2. Part of the software. 47 pp. - US EPA (2012). User's Guide for T.E.S.T. (version 4.1) (Toxicity Estimation Software Tool). Part of the software. 69 pp. - Zhu, H., Tropsha, A., Fourches, D., Varnek, A., Papa, E., Gramatica, P., Öberg, T., Dao, P., Cherkasov, A., and Tetko, I. V. 2008. Combinational QSAR Model of Chemical Toxicants Tested against *Tetrahymena pyriformis*. Journal of Chemical Information and Modeling 48:766 - 784.
9.3	Supporting information	-

1.4 QMRF: BCF baseline model v.02.07 (OASIS Catalogic v5.11.13)

1.0	QSAR identifier	
1.1	QSAR identifier (title)	BCF base-line model v.02.07
1.2	Other related models	-

1.3	Software coding the model	OASIS Catalogic v.5.11.13 [BCF base-line model v.02.07]; POPs v2.60.2 [BCF base-line model v.02.07]; Canadian POPs v1.2.3 [BCF base-line model v.02.07] http://oasis-lmc.org Laboratory of Mathematical Chemistry, University "Prof. Assen Zlatarov", 1 Yakimov Str. Burgas 8010, BULGARIA
2.0	General information	
2.1	Date of QMRF	10 March 2010
2.2	QMRF author and contact details	Laboratory of Mathematical Chemistry, University "Prof. Assen Zlatarov, " 1 Yakimov Str., Burgas 8010, BULGARIA http://www.oasis-lmc.org
2.3	Date of QMRF update(s)	02 December 2013
2.4	QMRF update(s)	-
2.5	Model developer(s) and contact details	S. Dimitrov, N. Dimitrova, D. Georgieva, T. Parkerton, M.Comber, M. Bonnell, O.Mekenyan. sdimitrov@btu.bg; ndimitrova@btu.bg; denitsa_georgieva@btu.bg; omekenya@btu.bg
2.6	Date of model development and/or publication	2005 December
2.7	References to main scientific papers and/or software package	S. Dimitrov, N. Dimitrova, T. Parkerton, M.Comber, M. Bonnell, O.Mekenyan. Base-line model for identifying the bioaccumulation potential of chemicals. SAR QSAR Environ Res, 16(6), 531-554, (2005). S. Dimitrov, G. Dimitrov, T. Pavlov, N. Dimitrova, G. Patlewiez, J. Niemela, O. Mekenyan. A stepwise Approach for defining the applicability domain of SAR and QSAR models. J Chem Inf Model, 45(4), 839 849, (2005).
2.8	Availability of information about the model	http://oasis-lmc.org/products/models/environmental-fate-and-ecotoxicity/bcf-base-line-model-(1).aspx
2.9	Availability of another QMRF for exactly the same model	-
3.0	Defining the endpoint – OECD Principle 1	
3.1	Species	Cyprinos carpio; salmonids
3.2	Endpoint	Environmental fate: BCF
3.3	Comment on the endpoint	BCF base-line model predicts bioconcentration factor (BCF, l/kg wet) in fish. Model accounts for a number of mitigating factors, such as molecular size, metabolism of parent chemical, water solubility and ionization.
3.4	Endpoint units	l/kg.weight
3.5	Dependent variable	log BCF
3.6	Experimental protocol	OECD 305

3.7	Endpoint data quality	High quality, chemicals provided by MITI (NITE), Japan; ExxonMobil
4.0	Defining the algorithm – OECD Principle 2	
4.1	Type of model	QSAR
4.2	Explicit algorithm	Prediction of BCF: The base-line concept for modeling the bioconcentration of chemicals is based on a reference curve delineating the maximum bioconcentration driven by hydrophobicity of chemicals (log BCFmax). Mitigating phenomena and chemical properties that can reduce bioconcentration potential, such as molecular size and flexibility, ionization, biotransformation, etc., are used as reducing factors of the maximum bioconcentration determined via the base-line. Parameterization of metabolism required the development of a fish liver simulator, given the shortage of fish metabolism data rat liver was used as an appropriate surrogate. 433 observed metabolism maps and expert knowledge were used to develop the metabolism simulator. The metabolism simulator consists of 497 transformations, of which 447 phase I and 50 phase II reactions. Non-linear least square method was used to estimate the model parameters.
4.3	Descriptors in the model	log Kow, metabolism, molecular size, ionization, water solubility.
4.4	Descriptor selection	-
4.5	Algorithm and descriptor generation	Not applicable
4.6	Software name and version for descriptor generation	Not applicable
4.7	Descriptor/Chemicals ratio	Not applicable
5.0	Defining the applicability domain – OECD Principle 3	
5.1	Description of the applicability domain of the model	The stepwise approach [6] was used to define the applicability domain of the model. It consists of the following sub-domain levels: - General parametric requirements – includes ranges of variation log Kow and MW, - Structural domain – based on atom-centered fragments (ACFs), - Mechanistic domain – identifies the mode of bioaccumulation of chemicals (partitioning in the organism lipids or binding to proteins). A chemical is considered In Domain if its log Kow and MW are within the specified ranges, its ACFs are presented in the training chemicals and if the mode of bioaccumulation is driven by the lipophilicity only. The information implemented in the applicability domain is extracted from the correctly predicted training chemicals used to build the model and in this respect, the applicability domain determines practically the interpolation space of the model.
5.2	Method used to	-

	assess the applicability domain	
5.3	Software name and version for applicability domain assessment	Domain Manager, Laboratory of Mathematical Chemistry University, "Prof. Assen Zlatarov", 1 Yakimov Str., Burgas 8010, BULGARIA
5.4	Limits of applicability	In order to belong to the model domain a target structure must meet the requirements of all the domain layers. -log Kow: Min -4.05 Max 16.07 -Molecular Weight: Min 16.04 Max 1131.21 -Water Solubility: Min 0 Max 1000000.06
6.0	Internal validation – OECD Principle 4	
6.1	Availability of the training set	Yes
6.2	Available information for the training set	CAS: Yes Chemical Name: Yes SMILES: Yes Formula: Yes INChI: No MOL file: No
6.3	Data for each descriptor variable for the training set	Yes
6.4	Data for the dependent variable (response) for the training set	Yes
6.5	Other information about the training set	The training set of the model consists of 705 chemicals and is a compilation of three databases: - 393 chemicals extracted from Biodegradation and Bioaccumulation Data of Existing Chemicals Based on the CSCL Japan (MITI database) [1]. - 167 chemicals tested by National Institute of Technology and Evaluation of Japan (NITE) using the same fish (*Cyprinus carpio*) [2]. - 145 BCF values extrapolated from dietary bioaccumulation experiments with salmonids [3]. MITI and NITE BCF data derived at the lowest concentration exposure have been used in the model development. All experimental data meet the OECD 305 protocol criteria and were generated based on the concentration of the parent chemicals only and not on the total amount of parent and metabolites (e.g., the total radioactivity). Another training database of documented fish and rat liver transformation maps for 433 organic compounds and expert knowledge was used to determine the principal transformations and to train the system to simulate the fish liver metabolism chemicals. The documented pathways were collected from scientific papers, monographs and databases accessible over the Internet.

6.6	Pre-processing of data before modelling	-
6.7	Statistics for goodness-of-fit	Statistics of the model: - $R^2 = 0.85$ - False negatives – 11 chemicals - False positive – 3 chemicals - Specificity (correct predicted not bioaccumulation chemicals/total not bioaccumulation chemicals) = 99% - Sensitivity (correct predicted bioaccumulation chemicals /total bioaccumulation chemicals) = 84%
6.8	Robustness – Statistics obtained by leave-one-outcross-validation	Not applicable
6.9	Robustness – Statistics obtained by leave-many-outcross-validation	Not applicable
6.10	Robustness – Statistics obtained by Y-scrambling	Not applicable
6.11	Robustness – Statistics obtained by bootstrap	Not applicable
6.12	Robustness – Statistics obtained by other methods	Not applicable
7.0	External validation - OECD Principle 4	
7.1	Availability of the external validation set	Yes
7.2	Available information for the external validation set	See 6.2
7.3	Data for each descriptor variable for external validation set	See 6.3
7.4	Data for the dependent variable for the external validation set	See 6.4
7.5	Other information about the external validation set	- The predictability of the model was evaluated on the basis of an external validation set of 176 chemicals provided by National Institute for Technology and Evaluation (NITE) Japan. The correctness of prediction for 59 chemicals identified to belong to the model applicability domain was 80%. For the rest of 117

		chemicals which do not belong to model applicability domain correctness of predictions was 50%.
7.6	Experimental design of test set	-
7.7	Predictivity – Statistics obtained by external validation	-
7.8	Predictivity - Assessment of the external validation set	-
7.9	Comments on the external validation of the model	-
8.0	interpretation Providing a mechanistic interpretation - OECD Principle 5	
8.1	Mechanistic basis of the model	The BCF base-line model consists of two major components: a model for predicting the maximum potential for bioaccumulation (log BCF_{max}) based solely on chemicals' lipophilicity and a set of mitigating factors that account for the reduction of the bioaccumulation potential of chemicals based on chemical (molecular size, ionization and water solubility) and organism (metabolism) dependent factors. Mathematical formulation of the model is: log BCF = log ($Pi_i(F_i(Kow_n/ (aKow +)^{2n})$)+ F_w* F_{ws}) where Kow is octanol-water partition coefficient, F_i stands for the set of mitigating factors: metabolism, molecular size, ionization, F_{ws} is water solubility factor, F_w is the organism water content. Further details on the mathematical formalism of the model can be reviewed in [4, 5]
8.2	A priori or a posteriori mechanistic interpretation	-
8.3	Other information about the mechanistic interpretation	-
9.0	Miscellaneous information	
9.1	Comments	-
9.2	Bibliography	S. Dimitrov, N. Dimitrova, T. Parkerton, M.Comber, M. Bonnell, O.Mekenyan. Base-line model for identifying the bioaccumulation potential of chemicals. SAR QSAR Environ Res, 16(6), 531-554, (2005). S. Dimitrov, G. Dimitrov, T. Pavlov, N. Dimitrova, G. Patlewiez, J. Niemela, O. Mekenyan. A stepwise Approach for defining the applicability domain of SAR and QSAR models. J Chem Inf Model, 45(4), 839 849, (2005). Chemicals Inspection and Testing Institute, Biodegradation and Bioaccumulation data of existing chemicals based on the CSCL

		Japan, Chemical Industry Ecology-Toxicology & Information Center, Japan, 1992, ISBN 4-98074-101-1. NITE, Biodegradation and Bioconcentration of the Existing Chemical Substances under the Chemical Substances Control Law, http://www.safe.nite.go.jp/english/db.html T. Parkerton. Phase II Report. The bioaccumulation of petroleum substances and their constituent hydrocarbons on the Canadian Designated Substances List (DSL), Exxon Mobil Biomedical Sciences Inc., 2004. S. Dimitrov, N. Dimitrova, D. Georgieva, K. Vasilev, T. Hatfield, J. Straka, and O. Mekenyan, SAR QSAR Environ. Res. 23, 2011,17–36
9.3	Supporting information	-

1.5 QMRF: Comparative analysis of estimated and measured BCF data (OECD 305; Müller & Nendza, 2011)

1.0	QSAR identifier	
1.1	QSAR identifier (title)	Comparative analysis of estimated and measured BCF data (OECD 305; Müller & Nendza, 2011): 13 QSARs for the estimation of the BCF based on log Kow
1.2	Other related models	-
1.3	Software coding the model	Not applicable; an Excel workbook is available which calculates the BCF for the 13 models.
2.0	General information	
2.1	Date of QMRF	04 Nov. 2013
2.2	QMRF author and contact details	BASF SE, Department of Product Safety, Ludwigshafen, Germany
2.3	Date of QMRF update(s)	-
2.4	QMRF update(s)	-
2.5	Model developer(s) and contact details	1) Veith et al. (1979) 2) Connell and Hawker (1988) 3) European Communities (2003) 4) Nendza (1991) 5) Mackay (1982) 6) Veith et al. (1983) 7) Bintein et al. (1993) 8) Schüürmann and Klein (1988) 9) Könemann and van Leeuwen (1980) 10) Lu et al. (1999) 11) Escuder-Gilabert et al. (2001) 12) Neely et al. (1974) 13) Zok et al. (1991)
2.6	Date of model development and/or	1) Veith et al. (1979) 2) Connell and Hawker (1988)

	publication	3) European Communities (2003) 4) Nendza (1991) 5) Mackay (1982) 6) Veith et al. (1983) 7) Bintein et al. (1993) 8) Schüürmann and Klein (1988) 9) Könemann and van Leeuwen (1980) 10) Lu et al. (1999) 11) Escuder-Gilabert et al. (2001) 12) Neely et al. (1974) 13) Zok et al. (1991)
2.7	References to main scientific papers and/or software package	Models evaluated in: Müller M, Nendza M (2011). Comparative analysis of estimated and measured BCF data (OECD 305). Federal Environment Agency (Umweltbundesamt), Texte 15/2011, Report no. UBA-FB 001435/E . 54 pp. References to the models: - Bintein S, Devillers J, Karcher W. 1993. Nonlinear Dependence of Fish Bioconcentration on n-Octanol/Water Partition Coefficient. SAR QSAR Environ. Res. 1: 29-39. - Connell DW, Hawker DW. 1988. Use of Polynomial Expressions to describe the Bioconcentration of Hydrophobic Chemicals by Fish. Ecotox. Environ. Saf. 16: 242-257. - Escuder-Gilabert L, Martin-Biosca Y, Sagrado S, Villanueva-Camanas RM, Medina-Hernandez MJ. 2001. Biopartitioning Micellar Chromatography to Predict Ecotoxicity. Analytica Chimica Acta 448: 173-185. - European Communities. 2003. Technical guidance document on risk assessment in support of Commission Directive 93/67/EEC on risk assessment for new notified substances, Commission Regulation (EC) No 1488/94 on risk assessment for existing substances, Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market. Joint Research Centre, Ispra, Italy: European Commission. - Köneman H, van Leeuwen K. 1980. Toxicokinetics in Fish: Accumulation and Elimination of Six Chlorobenzenes by Guppies. Chemosphere 9: 3-19. - Lu XX, Tao S, Cao J, Dawson RW. 1999. Prediction of Fish Bioconcentration Factors of Nonpolar Organic Pollutants based on Connectivity Indices. Chemosphere 39: 987-999. - Mackay D. 1982. Correlation of Bioconcentration Factors. Environ. Sci. Technol. 16: 274-278. - Neely WB, Branson DR, Blau GE. 1974. Partition Coefficients to Measure Bioconcentration Potential of Organic Chemicals in Fish. Env. Sci. Technol. 8: 1113-1115. - Nendza M. 1991. QSARs of bioconcentration: validity assessment of log Pow/log BCF correlations. In Bioaccumulation in aquatic systems, ed. Nagel, R. and Loskill, R. 43-66. Weinheim: VCH.

		- Schüürmann G, Klein W. 1988. Advances in Bioconcentration Prediction. Chemosphere 17: 1551-1574. - Veith GD, Defoe DL, and Bergstedt BV. 1979. Measuring and estimating the bioconcentration factor of chemicals in fish. J.Fish.Board Can. 36: 1040-1048. - Veith GD, Kosian P. 1983. Estimating Bioconcentration Potential from Octanol/Water Partition Coefficients. In: Physical Behaviour of PCBs in the Great Lakes. Mackay D, Paterson S, Eisenreich SJ, Simmons MS (Eds.), Ann Arbor Science Publishers, Ann Arbor, MI, U.S.A. - Zok S, Görge G, Kalsch W, Nagel R. 1991. Bioconcentration, Metabolism, and Toxicity of Substituted Anilines in the Zebrafish (Brachydanio rerio). Sci.Tot. Environ. 109/110: 411-421.
2.8	Availability of information about the model	The models are described and evaluated in: Müller M, Nendza M (2011). Comparative analysis of estimated and measured BCF data (OECD 305). Federal Environment Agency (Umweltbundesamt), Texte 15/2011, Report no. UBA-FB 001435/E . 54 pp.
2.9	Availability of another QMRF for exactly the same model	No (http://qsardb.jrc.it/qmrf/).
3.0	Defining the endpoint	
3.1	Species	Bioaccumulation potential estimated for fish
3.2	Endpoint	Bioconcentration factor (BCF)
3.3	Comment on the endpoint	Regulation (EC) No 1907/2006 [REACH], Annex 1X, 9.3.2 Bioaccumulation in aquatic species, preferably fish
3.4	Endpoint units	Bioconcentration factor (BCF): L/kg wet weight
3.5	Dependent variable	Bioconcentration factor (log BCF)
3.6	Experimental protocol	The bioconcentration of a substance can be determined according to OECD guideline 305.
3.7	Endpoint data quality	The test dataset used to develop the models vary in size from 6 to 154 compounds. Some models are based on rather heterogeneous datasets, while others are based on singe chemical classes (e. g. substituted anilines; see also 5.1).
4.0	Defining the algorithm	
4.1	Type of model	QSAR
4.2	Explicit algorithm	Model no. 1: log BCF = 0.85*logKow-0.7 Model no. 2: log BCF = 0.0069*POTENZ(logKow;4)-0.185*POTENZ(logKow;3)+1.55*POTENZ(logKow;2)-4.18*logKow+4.79 Model no. 3: log BCF = -0.2*POTENZ(logKow;2) + 2.74*logKow-4.72 Model no. 4: log BCF = 0.99*logKow-1.47*LOG(0.0000000497*POTENZ(10;logKow)+1;10)+0.0135 Model no. 5: log BCF = logKow-1.32 Model no. 6: log BCF = 0.79*logKow-0.4

		Model no. 7: log BCF = 0.91*logKow-1.975* LOG(0.00000068 *POTENZ(10;logKow)+1;10)-0.786 Model no. 8: log BCF = 0.75*logKow-0.32 Model no. 9: log BCF = 3.41*logKow-0.264*POTENZ(logKow;2)-5.513 Model no. 10: log BCF = 0.9*logKow-0.8 Model no. 11: log BCF = 0.74*logKow+0.8 Model no. 12: log BCF = 0.54*logKow+0.12 Model no. 13: log BCF = 0.67*logKow-0.18
4.3	Descriptors in the model	Log Kow
4.4	Descriptor selection	-
4.5	Algorithm and descriptor generation	Log Kow entered by user.
4.6	Software name and version for descriptor generation	-
4.7	Descriptor/Chemicals ratio	Descriptors: 1 Chemicals: 6 to 154, depending on model
5.0	Defining the applicability domain	
5.1	Description of the applicability domain of the model	The applicability domain is defined by the range of the log Kow of the training dataset. In some cases a recommended range is given for the log Kow. Some models are restricted to certain chemical classes based on the training dataset. In general, linear models give a fair approximation for the BCF for organic chemicals that are non-ionic, are not or very slowly metabolised and have a log Kow in the range of 1 to 6 (Pavan et al. 2006). This restriction applies to the following models: 1, 5, 6, 8, and 10 to 13. Model no. 1: heterogeneous dataset (*Pimephales promelas*); n = 55; r = 0.95 Model no. 2: heterogeneous dataset (fish (various)); n = 45 Model no. 3: heterogeneous dataset (fish (various)); n = 43; r = 0.883 Model no. 4: heterogeneous dataset (fish (various)); n = 132; model not derived by regression; therefore no statistical data available Model no. 5: heterogeneous dataset, mainly chlorinated hydrocarbons (fish (various)); n = 44; r = 0.95; s = 0.25 Model no. 6: heterogeneous dataset, mainly halogenated compounds (fish (various)); n = 122; r = 0.927; s = 0.49 Model no. 7: heterogeneous dataset (fish (various)); n = 154; r = 0.95; s = 0.347 Model no. 8: heterogeneous dataset, mainly chlorinated and polycyclic hydrocarbons (fish (various)); n = 32; r = 0.87; s = 0.54 Model no. 9: chlorobenzenes (*Poecilia reticulata*); n = 6; r = 0.999; s = 0.039

		Model no. 10: diverse non-polar chemicals (various fish); n = 80; r = 0.944 Model no. 11: diverse (various fish); n = 66; r = 0.917 Model no. 12: halogenated aromatics (*Salmo gairdneri*); n = 8; r = 0.949 Model no. 13: substituted anilines (*Brachydanio rerio*); n = 9; r = 0.934
5.2	Method used to assess the applicability domain	Log Kow and chemical class based on training dataset.
5.3	Software name and version for applicability domain assessment	-
5.4	Limits of applicability	-
6.0	Defining goodness-of-fit and robustness	
6.1	Availability of the training set	The complete datasets used to train the SAR equations used by the HYDROWIN program are available in the On-Line Help File of HYDROWIN v2.00.
6.2	Available information for the training set	In case of esters, information available on the fragments, the experimental and the estimated Kb (L/(mol*s). In case of other chemical classes, information on chemical name, CAS number and half-life data and corresponding pH available.
6.3	Data for each descriptor variable for the training set	The fragment substituent values which are used to calculate the hydrolysis rate constant are listed in Appendix E.
6.4	Data for the dependent variable (response) for the training set	See 6.2
6.5	Other information about the training set	-
6.6	Pre-processing of data before modelling	-
6.7	Statistics for goodness-of-fit	See 5.1
7.0	Defining predictivity	
7.1	Availability of the external validation set	Not available
7.2	Available information for the external validation set	-

7.3	Data for each descriptor variable for external validation set	-
7.4	Data for the dependent variable for the external validation set	-
7.5	Other information about the external validation set	-
7.6	Experimental design of test set	-
7.7	Predictivity – Statistics obtained by external validation	-
7.8	Predictivity - Assessment of the external validation set	-
8.0	Providing a mechanistic interpretation	
8.1	Mechanistic basis of the model	Quantitative structure-activity relationships (QSAR) make use of the fact that bioaccumulation of stable organic compounds is governed by partitioning between aqueous and lipid phases. The predominant process of passive diffusion is frequently formalized in log Kow-dependent QSAR models. It is often assumed, that the log Kow-based BCF estimates represent a 'worst case' reference point. Estimating bioconcentration factors (BCF) from octanol/water partition coefficients (log Kow) is well established and essentially valid for neutral organics of intermediate lipophilicity (0 < log KOW < 6) (European Communities, 2003; Nendza, 1991; Nendza, 1998; Dearden, 2004).
9.0	Miscellaneous information	
9.1	Comments	-
9.2	Bibliography	- Bintein S, Devillers J, Karcher W. 1993. Nonlinear Dependence of Fish Bioconcentration on n-Octanol/Water Partition Coefficient. SAR QSAR Environ. Res. 1: 29-39. - Connell DW, Hawker DW. 1988. Use of Polynomial Expressions to describe the Bioconcentration of Hydrophobic Chemicals by Fish. Ecotox. Environ. Saf. 16: 242-257. - Dearden JC. 2004. QSAR modeling of bioaccumulation. In Predicting chemical toxicity and fate, ed. Cronin, M. T. D. and Livingstone, D. J. 333-55. Boca Raton: CRC Press. - Escuder-Gilabert L, Martin-Biosca Y, Sagrado S, Villanueva-Camanas RM, Medina-Hernandez MJ. 2001. Biopartitioning Micellar Chromatography to Predict Ecotoxicity. Analytica Chimica Acta 448: 173-185. - European Communities. 2003. Technical guidance document on

		risk assessment in support of Commission Directive 93/67/EEC on risk assessment for new notified substances, Commission Regulation (EC) No 1488/94 on risk assessment for existing substances, Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market. Joint Research Centre, Ispra, Italy: European Commission. - Köneman H, van Leeuwen K. 1980. Toxicokinetics in Fish: Accumulation and Elimination of Six Chlorobenzenes by Guppies. Chemosphere 9: 3-19. - Lu XX, Tao S, Cao J, Dawson RW. 1999. Prediction of Fish Bioconcentration Factors of Nonpolar Organic Pollutants based on Connectivity Indices. Chemosphere 39: 987-999. - Mackay D. 1982. Correlation of Bioconcentration Factors. Environ. Sci. Technol. 16: 274-278. - Müller M, Nendza M (2011). Comparative analysis of estimated and measured BCF data (OECD 305). Federal Environment Agency (Umweltbundesamt), Texte 15/2011, Report no. UBA-FB 001435/E . 54 pp. - Neely WB, Branson DR, Blau GE. 1974. Partition Coefficients to Measure Bioconcentration Potential of Organic Chemicals in Fish. Env. Sci. Technol. 8: 1113-1115. - Nendza M. 1991. QSARs of bioconcentration: validity assessment of log Pow/log BCF correlations. In Bioaccumulation in aquatic systems, ed. Nagel, R. and Loskill, R. 43-66. Weinheim: VCH. - Nendza M. 1998. Structure-activity relationships in environmental sciences. London, Great Britain: Chapman & Hall. - Pavan M, Woth AP, Netzeva TI (2006). Review of QSAR models for bioaconcentration. EUR 22327EN, European Commission – Joint Reasearch Centre, 123 pp. - Schüürmann G, Klein W. 1988. Advances in Bioconcentration Prediction. Chemosphere 17: 1551-1574. - Veith GD, Defoe DL, and Bergstedt BV. 1979. Measuring and estimating the bioconcentration factor of chemicals in fish. J.Fish.Board Can. 36: 1040-1048. - Veith GD, Kosian P. 1983. Estimating Bioconcentration Potential from Octanol/Water Partition Coefficients. In: Physical Behaviour of PCBs in the Great Lakes. Mackay D, Paterson S, Eisenreich SJ, Simmons MS (Eds.), Ann Arbor Science Publishers, Ann Arbor, MI, U.S.A. - Zok S, Görge G, Kalsch W, Nagel R. 1991. Bioconcentration, Metabolism, and Toxicity of Substituted Anilines in the Zebrafish (Brachydanio rerio). Sci.Tot. Environ. 109/110: 411-421.

9 ANNEX 2: QPRF'S: CRITERIA FOR THE APPLICABILITY DOMAIN

The information if the substance meets the criteria of the applied (Q)SAR models' applicability domains is given according to the (Q)SAR Prediction Reporting Format (QPRF) following the OECD principles stated in REACH Guidance R.6 (ECHA, 2008).

9.1 QPRF: BCFBAF v3.01 (EPI Suite v4.11)

1.	Substance	CAS 103597-45-1	
2.	General information		
2.1	Date of QPRF	22 Sep. 2014	
2.2	QPRF author and contact details	BASF SE, Dept. for Product Safety, Ludwigshafen, Germany	
3.	Prediction		
3.1	Endpoint (OECD Principle 1)	Endpoint	Bioaccumulation (aquatic)
		Dependent variable	- Bioconcentration factor (BCF) - Bioaccumulation factor (BAF; 15 °C) - Biotransformation rate (kM) and half-life
3.2	Algorithm (OECD Principle 2)	Model or submodel name	BCFBAF Submodels: 1) Bioconcentration factor (BCF; Meylan et al., 1997/1999) 2) Biotransformation rate in fish (kM; Arnot et al., 2008a/b) 3) Arnot & Gobas BAF and steady-state BCF Arnot & Gobas, 2003)
		Model version	v. 3.01
		Reference to QMRF	Estimation of Bioconcentration, bioaccumulation and biotransformation in fish using BCFBAF v3.01 (EPI Suite v4.11)
		Predicted value (model result)	See Table 14
		Input for prediction	Chemical structure via CAS number or SMILES; log Kow (optional)
		Descriptor values	- SMILES: structure of the compound as SMILES notation - log Kow - Molecular weight
3.3	Applicability domain (OECD principle 3)	Domains:	
		1) Bioconcentration factor (BCF; Meylan et al., 1997/1999)	
		a) Ionic/non-Ionic	The substance is ionic (pKa = 7, phenolic group, but according to the very poor water solubility this is not expected to have a

			significant effect on the substances behaviour under environmentally relevant conditions).
		b) Molecular weight (range of test data set): - Ionic: 68.08 to 991.80 - Non-ionic: 68.08 to 959.17 (On-Line BCFBAF Help File, Ch. 7.1.3 Estimation Domain and Appendix G)	The substance is within range (659 g/mol).
		c) log Kow (range of test data set): - Ionic: -6.50 to 11.26 - Non-ionic: -1.37 to 11.26 (On-Line BCFBAF Help File, Ch. 7.1.3 Estimation Domain and Appendix G)	The substance is not within range (log Kow = 12.46).
		d) Maximum number of instances of correction factor in any of the training set compounds (On-Line BCFBAF Help File, Appendix E)	Not exceeded.
		2) Biotransformation rate in fish (kM; Arnot et al., 2008a/b)	
		a) The substance does not appreciably ionize at physiological pH. (On-Line BCFBAF Help File, Ch. 7.2.3)	Fulfilled
		b) Molecular weight (range of test data set): 68.08 to 959.17 (On-Line BCFBAF Help File, Ch. 7.2.3)	The substance is within range (659 g/mol).
		c) The molecular weight is ≤ 600 g/mol. (On-Line BCFBAF Help File, Ch. 7.2.3)	Not fulfilled
		d) Log Kow: 0.31 to 8.70 (On-Line BCFBAF Help File, Ch. 7.2.3)	The substance is not within range (log Kow = 12.46).
		e) The substance is no metal or organometal, pigment or dye, or a perfluorinated substance. (On-Line BCFBAF Help File, Ch. 7.2.3)	Fulfilled
		f) Maximum number of instances of biotransformation fragments	Exceeded. Fragment "number of fused 5 -carbon aromatic rings" was identified by the model but no

		in any of the training set compounds (On-Line BCFBAF Help File, Appendix F)	coeficient was assigned.
		3) Arnot & Gobas BAF and steady-state BCF Arnot & Gobas, 2003)	
		a) Log Kow ≤ 9 (On-Line BCFBAF Help File, Ch. 7.3.1)	Not fulfilled
		b) The substance does not appreciably ionize. (On-Line BCFBAF Help File, Ch. 7.3.1)	Fulfilled (pKa = 7, phenolic group, but according to the very poor water solubility this is not expected to have a significant effect on the substances behaviour under environmentally relevant conditions).
		c) The substance is no pigment, dye, or perfluorinated substance. (On-Line BCFBAF Help File, Ch. 7.3.1)	Fulfilled
3.4	The uncertainty of the prediction (OECD principle 4)	1. Bioconcentration factor (BCF; Meylan et al., 1997/1999) Statistical accuracy of the training data set (non-ionic plus ionic data): - Correlation coefficient (r^2) = 0.833 - Standard deviation = 0.502 log units - Absolute mean error = 0.382 log units 2. Biotransformation Rate in Fish (kM) Statistical accuracy (training set): - Correlation coefficient (r^2) = 0.821 - Correlation coefficient (Q^2) = 0.753 - Standard deviation = 0.494 log units - Absolute mean error = 0.383 log units 3. Arnot-Gobas BAF/BCF model No information on the statistical accuracy given in the documentation.	
3.5	The chemical mechanisms according to the model underpinning the predicted result (OECD principle 5)	1. The BCF model is mainly based on the relationship between bioconcentration and hydrophobicity. The model also takes into account the chemical structure and the ionic/non-ionic character of the substance. 2. Bioaccumulation is the net result of relative rates of chemical inputs to an organism from multimedia exposures (e.g., air, food, and water) and chemical outputs (or elimination) from the organism. 3. The model includes mechanistic processes for bioconcentration	

		and bioaccumulation such as chemical uptake from the water at the gill surface (BCFs and BAFs) and the diet (BAFs only), and chemical elimination at the gill surface, fecal egestion, growth dilution and metabolic biotransformation (Arnot and Gobas 2003). Other processes included in the calculations are bioavailability in the water column (only the freely dissolved fraction can bioconcentrate) and absorption efficiencies at the gill and in the gastrointestinal tract.

References

- Arnot JA, Gobas FAPC. 2003. A generic QSAR for assessing the bioaccumulation potential of organic chemicals in aquatic food webs. QSAR and Combinatorial Science 22: 337-345.
- Arnot JA, Mackay D, Parkerton TF, Bonnell M. 2008a. A database of fish biotransformation rates for organic chemicals. Environmental Toxicology and Chemistry 27(11), 2263-2270.
- Arnot JA, Mackay D, Bonnell M. 2008b.Estimating metabolic biotransformation rates in fish from laboratory data. Environmental Toxicology and Chemistry 27: 341-351.
- Meylan, W.M., Howard, P.H, Aronson, D., Printup, H. and S. Gouchie. 1997. "Improved Method for Estimating Bioconcentration Factor (BCF) from Octanol-Water Partition Coefficient", SRC TR-97-006 (2nd Update), July 22, 1997; prepared for: Robert S. Boethling, EPA-OPPT, Washington, DC; Contract No. 68-D5-0012; prepared by: ; Syracuse Research Corp., Environmental Science Center, 6225 Running Ridge Road, North Syracuse, NY 13212.
- Meylan, WM, Howard, PH, Boethling, RS et al. 1999. Improved Method for Estimating Bioconcentration / Bioaccumulation Factor from Octanol/Water Partition Coefficient. Environ. Toxicol. Chem. 18(4): 664-672 (1999).
- US EPA (2012). On-Line BCFBAF Help File.

Identified Correction Factors (Appendix E), Biotransformation Fragments and Coefficient values (Appendix F)

Appendix E: not applicable, no corrections factors used

Appendix F

The Training Set used to derive the Coefficient Values listed below contained a total of 421 compounds (see Appendix I for the compound list).

Fragment Description	Coefficient value	No. compounds containing fragment in total training set	Maximum number of each fragment in any individual compound	No. of instances of each fragment for the current substance
Aromatic alcohol [-OH]	-0.47273947	26	2	2
Carbon with 4 single bonds & no hydrogens	-0.29842827	47	10	4
Alkyl substituent on aromatic ring	0.17805958	88	6	1
Triazole Ring	0.32253333	4	1	2
Aromatic-CH2	-0.33650743	30	4	1
Aromatic-H	0.26637806	305	15	12
Methyl [-CH3]	0.24510529	170	12	10
-CH2- [linear]	0.02418707	109	28	2
Number of fused 6-carbon aromatic rings	-0.577854	67	5	2

Benzene	-0.427728	197	3	2

Assessment of the Applicability Domain Based on Molecular Weight and log Kow
1. Bioconcentration Factor (BCF; Meylan et al., 1997/1999)

Training set: Molecular weights	Ionic	Non-ionic
Minimum	68.08	68.08
Maximum	991.80	959.17
Average	244.00	244.00
Assessment of molecular weight	Molecular weight within range of training set.	
Training set: Log Kow	Ionic	Non-ionic
Minimum	-6.50	-1.37
Maximum	11.26	11.26
Assessment of log Kow	Log Kow outside of range of training set. Therefore, the estimate may be less accurate.	

2. Biotransformation Rate in Fish (kM; Arnot et al., 2008a/b)

Training set: Molecular weights	
Minimum	68.08
Maximum	959.17
Average	259.75
Assessment of molecular weight	Molecular weight within range of training set, but exceeds 600 g/mol. Therefore, the estimate may be less accurate.
Training set: Log Kow	
Minimum	0.31
Maximum	8.70
Assessment of log Kow	Log Kow outside of range of training set. Therefore, the estimate may be less accurate.

3. Arnot-Gobas BAF/BCF (Arnot & Gobas, 2003)

Assessment of log Kow	Log Kow > 9; therefore, the estimate may be highly uncertain.

9.2 VEGA v1.0.8: BCF models

9.2.1 QPRF: CAESAR v2.1.13 (VEGA v1.0.8)

The applicability domain of predictions is assessed using an Applicability Domain Index (ADI) that has values from 0 (worst case) to 1 (best case). The ADI is calculated by grouping several other indices, each one taking into account a particular issue of the applicability domain. Most of the indices are based on the calculation of the most similar compounds found in the training and test set of the model, calculated by a similarity index that consider molecule's fingerprint and structural aspects (count of atoms, rings and relevant fragments). Note that when the experimental value for the given compound is found, the applicability Domain indices are calculated only considering this value, without taking into account the first*n*similar compounds.
For each index, including the final ADI, three intervals for its values are defined, such that the first interval corresponds to a positive evaluation, the second one corresponds to a suspicious evaluation and the last one corresponds to a negative evaluation.

Following, all applicability domain components are reported along with their explanation and the intervals used. Furthermore, the specific index of the substance is given.

9.2.1.1 Similar molecules with known experimental value.

This index takes into account how similar are the first two most similar compounds found. Values near 1 mean that the predicted compound is well represented in the dataset used to build the model, otherwise the prediction could be an extrapolation.
Defined intervals are:

1 >= index > 0.9	strongly similar compounds with known experimental value in the training set have been found
0.9 >= index > 0.75	only moderately similar compounds with known experimental value in the training set have been found
index <= 0.75	no similar compounds with known experimental value in the training set have been found

The substance has a similarity index of 0.716.

9.2.1.2 Accuracy (average error) of prediction for similar molecules.

This index takes into account the error in prediction for the two most similar compounds found. Values near 0 mean that the predicted compounds falls in an area of the model's space where the model gives reliable predictions, otherwise the greater is the value, the worse the model behaves.
Defined intervals are:

index < 0.5	accuracy of prediction for similar molecules found in the training set is good
0.5 <= index <= 1.0	accuracy of prediction for similar molecules found in the training set is not optimal
index > 1.0	accuracy of prediction for similar molecules found in the training set is not adequate

The substance has an accuracy index of 0.555.

9.2.1.3 Concordance with similar molecules (average difference between target compound prediction and experimental values of similar molecules).

This index takes into account the difference between the predicted value and the experimental values of the two most similar compounds. Values near 0 mean that the prediction made agrees with the experimental values found in the model's space, thus the prediction is reliable.
Defined intervals are:

index < 0.5	similar molecules found in the training set have experimental values that agree with the target compound predicted value
0.5 <= index <= 1.0	similar molecules found in the training set have experimental values that slightly disagree with the target compound predicted value
index > 1.0	similar molecules found in the training set have experimental values that completely disagree with the target compound predicted value

The substance has a concordance index of 0.334.

9.2.1.4 Maximum error of prediction among similar molecules.

This index takes into account the maximum error in prediction among the two most similar compounds. Values near 0 means that the predicted compounds falls in an area of the model's space where the model gives reliable predictions without any outlier value.
Defined intervals are:

index < 0.5	the maximum error in prediction of similar molecules found in the training set has a low value, considering the experimental variability
0.5 <= index < 1.0	the maximum error in prediction of similar molecules found in the training set has a moderate value, considering the experimental variability
index >= 1.0	the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability

The substance has a max error index of 0.9.

9.2.1.5 Atom Centered Fragments similarity check.

This index takes into account the presence of one or more fragments that aren't found in the training set, or that are rare fragments. First order atom centered fragments from all molecules in the training set are calculated, then compared with the first order atom centered fragments from the predicted compound; then the index is calculated as following: a first index RARE takes into account rare fragments (those who occur less than three times in the training set), having value of 1 if no such fragments are found, 0.85 if up to 2 fragments are found, 0.7 if more than 2 fragments are found; a second index NOTFOUND takes into account not found fragments, having value of 1 if no such fragments are found, 0.6 if a fragments is found, 0.4 if more than 1 fragment is found. Then, the final index is given as the product RARE * NOTFOUND.
Defined intervals are:

index = 1	all atom centered fragment of the compound have been found in the compounds of the training set
1 > index >= 0.7	some atom centered fragment of the compound have not been found in the compounds of the training set or are rare fragments
index < 0.7	a prominent number of atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments

The substance has an ACF matching index of 0.7.

9.2.1.6 Descriptors noise sensitivity analysis.

This index checks whether the predicted compound falls in a reliable and stable descriptors space or not. A sequence of random scrambling (noise) is applied to the descriptors calculated for the considered compound, and it is checked if the perturbation of descriptors lead to a significant change in the prediction; if the studied descriptors space is stable, these changes should be of little entity. After a large number of such random scrambling, a final index is calculated.
Defined intervals are:

1 >= index > 0.8	predictions has a good response to noise scrambling, thus shows a good reliability
0.8 >= index > 0.5	predictions has a not so good response to noise scrambling, thus shows an uncertain reliability
index <= 0.5	predictions has a bad response to noise scrambling, thus shows a low reliability

The substance has a noise sensitivity of 0.937.

9.2.1.7 Model descriptors range check.

This index checks if the descriptors calculated for the predicted compound are inside the range of descriptors of the training and test set. The index has value 1 if all descriptors are inside the range, 0 if at least one descriptor is out of the range.
Defined intervals are:

index = 1	descriptors for this compound have values inside the descriptor range of the compounds of the training set
index = 0	descriptors for this compound have values outside the descriptor range of the compounds of the training set

The substance' descriptors range check is 0 (=false).

9.2.1.8 Global AD Index.

The final global index takes into account all the previous indices, in order to give a general global assessment on the applicability domain for the predicted compound.
Defined intervals are:

1 >= index > 0.85	predicted substance is into the Applicability Domain of the model
0.85 >= index > 0.75	predicted substance could be out of the Applicability Domain of the model
index <= 0.75	predicted substance is out of the Applicability Domain of the model

The substance has a global AD index of 0.

9.2.1.9 Detailed expert analysis

The result of the model may not be reliable. The following issues were noted by the model:
1) No similar compounds with known experimental value have been found in the training set.
2) Accuracy of prediction for similar molecules found in the training set is not optimal.
3) The maximum error in prediction of similar molecules found in the training set has a moderate value, considering the experimental variability.
4) Some atom centered fragments of the compound have not been found in the compounds of the training set or are rare fragments.
5) Descriptors for this compound have values outside the descriptor range of the compounds of the training set.
Regarding the complex structure of the substance, it is very likely that no similar compounds are available in the training set. Therefore, the reliability of the prediction may be low.

The model detected a structural alert which is listed and discussed in detail in the paragraph below.
Structural Alerts: Polar Groups: PG 06 = OH group
The substance contains two polar OH groups. The presence of polar groups increases hydrophilicity, related to lower values of BCF.
References:
VEGA Guide to BCF Model version 2.1.13 implemented in the VEGA tool v1.0.8

9.2.2 QPRF: BCF Read-Across v1.0.2 (VEGA v1.0.8)

The applicability domain of predictions is assessed using an Applicability Domain Index (ADI) that has values from 0 (worst case) to 1 (best case). The ADI is calculated by grouping several other

indices, each one taking into account a particular issue of the applicability domain. For each index, including the final ADI, two intervals for its values are defined, such that the first interval corresponds to a positive evaluation, and the second one corresponds to a negative evaluation.
Following, all applicability domain components are reported along with their explanation.
Furthermore, the specific index of the substance is given.

9.2.2.1 Highest similarity found for similar compounds.

This index takes into account the maximum value of similarity among the three most similar compounds found. Values higher than 0.7 mean that at least one compound with a good structural similarity with the chemical to be predicted has been found. Values lower than 0.7 mean that no remarkably similar compounds have been found, and the read-across could be not reliable.
Defined intervals are:

index >= 0.85	the highest similarity value found for similar compounds is adequate for a reliable read-across
index < 0.85	the highest similarity value found for similar compounds is not adequate for a reliable read-across

The substance has a maximum value of similarity of 0.766.

9.2.2.2 Lowest similarity found for similar compounds.

This index takes into account the minimum value of similarity among the three most similar compounds found. Values higher than 0.6 mean that also the least similar among the three compounds has an acceptable structural similarity with the chemical to be predicted. Values lower than 0.6 mean that the read-across could be not reliable.
Defined intervals are:

index >= 0.7	the lowest similarity value found for similar compounds is adequate for a reliable read-across
index < 0.7	the lowest similarity value found for similar compounds is not adequate for a reliable read-across

The substance has a minimum value of similarity of 0.701.

9.2.2.3 Global AD Index.

The final global index takes into account the previous indices, in order to give a general global assessment on the applicability domain for the predicted compound. If at least one of the previous indices has a negative evaluation, the final global index will result in an assessment of unreliability; if all indices have positive evaluation, then the global index will result in an assessment of reliability. In both cases, the global index value is calculated as the average value of the similarity index for the three compounds taken into account for the read-across.
The substance has a global AD index of 0.725.
Read-across seems to be unreliable due to low similarity in found molecules.

References:
VEGA Guide to BCF Read-Across version 1.0.2 implemented in the VEGA tool v1.0.8

9.2.3 QPRF: Meylan v1.0.2 (VEGA v1.0.8)

The applicability domain of predictions is assessed using an Applicability Domain Index (ADI) that has values from 0 (worst case) to 1 (best case). The ADI is calculated by grouping several other indices, each one taking into account a particular issue of the applicability domain. Most of the indices are based on the calculation of the most similar compounds found in the training and test set of the model, calculated by a similarity index that consider molecule's fingerprint and structural aspects (count of atoms, rings and relevant fragments). Note that when the experimental value for the given compound is found, the applicability Domain indices are calculated only considering this value, without taking into account the first *n* similar compounds.
For each index, including the final ADI, three intervals for its values are defined, such that the first interval corresponds to a positive evaluation, the second one corresponds to a suspicious evaluation and the last one corresponds to a negative evaluation.
Following, all applicability domain components are reported along with their explanation and the intervals used. Furthermore, the specific index of the substance is given.

9.2.3.1 Similar molecules with known experimental value.

This index takes into account how similar are the first two most similar compounds found. Values near 1 mean that the predicted compound is well represented in the dataset used to build the model, otherwise the prediction could be an extrapolation.
Defined intervals are:

1 >= index > 0.9	strongly similar compounds with known experimental value in the training set have been found
0.9 >= index > 0.75	only moderately similar compounds with known experimental value in the training set have been found
index <= 0.75	no similar compounds with known experimental value in the training set have been found

The substance has a similarity index of 0.761.

9.2.3.2 Accuracy (average error) of prediction for similar molecules.

This index takes into account the error in prediction for the two most similar compounds found. Values near 0 mean that the predicted compounds falls in an area of the model's space where the model gives reliable predictions, otherwise the greater is the value, the worse the model behaves.
Defined intervals are:

index < 0.5	accuracy of prediction for similar molecules found in the training set is good
0.5 <= index <= 1.0	accuracy of prediction for similar molecules found in the training set is not optimal
index > 1.0	accuracy of prediction for similar molecules found in the training set is not adequate

The substance has an accuracy index of 0.39.

9.2.3.3 Concordance with similar molecules (average difference between target compound prediction and experimental values of similar molecules).

This index takes into account the difference between the predicted value and the experimental values of the two most similar compounds. Values near 0 mean that the prediction made agrees with the experimental values found in the model's space, thus the prediction is reliable.
Defined intervals are:

index < 0.5	similar molecules found in the training set have experimental values that agree with the target compound predicted value
0.5 <= index <= 1.0	similar molecules found in the training set have experimental values that slightly disagree with the target compound predicted value
index > 1.0	similar molecules found in the training set have experimental values that completely disagree with the target compound predicted value

The substance has a concordance index of 1.616.

9.2.3.4 Maximum error of prediction among similar molecules.

This index takes into account the maximum error in prediction among the two most similar compounds. Values near 0 means that the predicted compounds fall in an area of the model's space where the model gives reliable predictions without any outlier value.
Defined intervals are:

index < 0.5	the maximum error in prediction of similar molecules found in the training set has a low value, considering the experimental variability
0.5 <= index < 1.0	the maximum error in prediction of similar molecules found in the training set has a moderate value, considering the experimental variability
index >= 1.0	the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability

The substance has a max error index of 0.72.

9.2.3.5 LogP reliability.

This index takes into account the reliability of the logP value used in the model. Note that the Meylan BCF model is strongly based on the logP prediction of the compound, thus this index is highly relevant for the assessment of the final prediction. The reliability of the logP value comes from the assessment of the VEGA LogP model (that provides the used logP value), which is also provided in the “Prediction summary” section of the report.
Defined intervals are:

index = 1	the maximum error in prediction of similar molecules found in the training set has a low value, considering the experimental variability
index = 0.7	the maximum error in prediction of similar molecules found in the training set has a moderate value, considering the experimental variability
index = 0	the maximum error in prediction of similar molecules found in the training set has a high value, considering the experimental variability

The substance has a LogP reliability index of 0.

9.2.3.6 Model descriptors range check.

This index checks if the descriptors calculated for the predicted compound are inside the range of descriptors of the training and test set. The index has value 1 if all descriptors are inside the range, 0 if at least one descriptor is out of the range.
Defined intervals are:

index = 1	descriptors for this compound have values inside the descriptor range of the compounds of the training set
index = 0	descriptors for this compound have values outside the descriptor range of the compounds of the training set

The substance' descriptors range check is 1 (= true).

9.2.3.7 Global AD Index.

The final global index takes into account all the previous indices, in order to give a general global assessment on the applicability domain for the predicted compound.
Defined intervals are:

1 >= index > 0.85	predicted substance is into the Applicability Domain of the model
0.85 >= index > 0.75	predicted substance could be out of the Applicability Domain of the model
index <= 0.75	predicted substance is out of the the Applicability Domain of the model

The substance has a global AD index of 0.75.

9.2.3.8 Detailed expert analysis

- only moderately similar compounds with known experimental value in the training set have been found
- similar molecules found in the training set have experimental values that strongly disagree with the target compound predicted value
- the maximum error in prediction of similar molecules found in the training set has a moderate value, considering the experimental
variability
- reliability of logP value used by the model is not adequate

Regarding the complex structure of the substance, it is very likely that no similar compounds are available in the training set. Therefore, the reliability of the prediction may be low.

References:
VEGA Guide to BCF Meylan Model version 1.0.2 implemented in the VEGA tool v1.0.8

9.3 US EPA T.E.S.T. v4.1: Bioaccumulation

QPRF: US EPA T.E.S.T. v4.1

1.	Substance	CAS 103597-45-1	
2.	General information		
2.1	Date of QPRF	23 Sep. 2014	
2.2	QPRF author and contact details	BASF SE; Dept. for Product Safety, Ludwigshafen, Germany	
3.	Prediction		
3.1	Endpoint (OECD Principle 1)	Endpoint	Bioaccumulation (aquatic)
		Dependent variable	Bioconcentration factor (BCF)
3.2	Algorithm (OECD Principle 2)	Model or submodel name	US EPA T.E.S.T. v4.1: 1) Hierarchical clustering 2) FDA method 3) Single model 4) Group contribution 5) Nearest neighbour

<table>
<tr><td></td><td></td><td></td><td>6) Consensus</td></tr>
<tr><td></td><td></td><td>Model version</td><td>v. 4.1</td></tr>
<tr><td></td><td></td><td>Reference to QMRF</td><td>Estimation of bioaccumulation in fish using T.E.S.T. v4.1</td></tr>
<tr><td></td><td></td><td>Predicted value (model result)</td><td>See Table 14</td></tr>
<tr><td></td><td></td><td>Input for prediction</td><td>Chemical structure via
CAS number, SMILES, MDL molfile, structure (drawing)</td></tr>
<tr><td></td><td></td><td>Descriptor values</td><td>Molecular descriptors (calculated by T.E.S.T.)</td></tr>
<tr><td>3.3</td><td>Applicability domain
(OECD principle 3)</td><td>General remarks</td><td>Predictions are considered only from valid models. Models which do not meet the constraints are listed in the output with a corresponding remark. If the substance is not within the applicability domain, no BCF is calculated.</td></tr>
<tr><td></td><td></td><td>Hierarchical clustering</td><td>In domain</td></tr>
<tr><td></td><td></td><td>FDA method</td><td>In domain</td></tr>
<tr><td></td><td></td><td>Single model</td><td>Not In domain</td></tr>
<tr><td></td><td></td><td>Group contribution</td><td>Not In domain</td></tr>
<tr><td></td><td></td><td>Nearest neighbour</td><td>In domain</td></tr>
<tr><td></td><td></td><td>Consensus</td><td>In domain</td></tr>
<tr><td>3.4</td><td>The uncertainty of the prediction
(OECD principle 4)</td><td colspan="2">The uncertainty of the predictions can be assessed by comparing the mean average error (MAE) of the entire dataset with the MAE of the dataset restricted to substances with a similarity coefficient (SC) of ≥ 0.5. If the MAE for the entire set is lower than the MAE for the similar substances (SC ≥ 0.5), the confidence in the predicted BCF value is high.
The table below lists the information on q^2(leave one out correlation coefficient), r^2(correlation coefficient), MAE and SC of the models.
Based on the MAE of the external and the training dataset, the confidence in the estimated BCF is assessed as follows.
<table>
<tr><td rowspan="2">Model</td><td colspan="2">Confidence in estimated BCF</td></tr>
<tr><td>External test set:</td><td>Training set:</td></tr>
<tr><td>Consensus method</td><td>low</td><td>low</td></tr>
<tr><td>Hierarchical clustering</td><td>low</td><td>low</td></tr>
<tr><td>Single model</td><td>N/A</td><td>N/A</td></tr>
<tr><td>Group contribution</td><td>N/A</td><td>N/A</td></tr>
<tr><td>FDA</td><td>low</td><td>low</td></tr>
<tr><td>Nearest neighbor</td><td>low</td><td>low</td></tr>
</table>
</td></tr>
</table>

3.5	The chemical mechanisms according to the model underpinning the predicted result (OECD principle 5)	Molecular descriptors are used to develop the models. The overall pool of descriptors in the software contain 797 2-dimensional descriptors of the following classes: E-state values and E-state counts, constitutional descriptors, topological descriptors, walk and path counts, connectivity, information content, 2d autocorrelation, Burden eigenvalue, molecular property (such as the octanol-water partition coefficient), Kappa, hydrogen bond acceptor/donor counts, molecular distance edge, and molecular fragment counts. The descriptors used to describe the compound can be viewed in the model output details.

Detailed information on q^2 (leave one out correlation coefficient), r^2 (correlation coefficient), MAE and SC:
Model details:

Method	Predicted value		Model statistics			MAE (in log10)			
						External test set		Training set	
	log BCF	BCF	r²	q²	No. of chemicals	Entire set	SC >= 0.5	Entire set	SC >= 0.5
Consensus method	2.01	101.85	-	-	-	0.51	0.76	0.42	0.64
Hierarchical clustering	3.22	1,666.21 (0.62-4510821.80)	0.662 - 0.764	0.569 - 0.705	114 - 118 (cluster models: 2)	0.54	0.90	0.23	0.37
Single model	N/A	N/A	0.764	0.733	540	0.54	N/A	0.53	N/A
Group contribution	N/A	N/A	0.719	0.527	499	0.62	N/A	0.60	N/A
FDA	0.99	9.71 (0.78-12.92)	0.906	0.665	40	0.57	0.82	0.53	1.22
Nearest neighbor	1.81	65.28	-	-	3	0.60	0.96	0.55	0.86

Legend:
SC = similarity coefficient
r² = correlation coefficient
q² = leave one out correlation coefficient

9.4 BCF baseline model v.02.07 (OASIS Catalogic v5.11.13)

MODEL DOMAIN
Parametric domain: In domain (100%)
- log Kow (range: -4.049 - 16.074): 12.7 (calculated)
- molecular weight (range: 16.041 - 1131.206 g/mol): 662.8766 g/mol
- water solubility (range: 0 - 1000000 mg/L): 0.000005 mg/L (< 5 ng/L, measured)

Structural domain: In domain (35%): 35% correct fragments, 0% incorrect fragments, 65% unknown fragments
Mechanistic domain: In domain (100%)
With regard to the parametric and the mechanistic domain, the test substance is within the applicability domain of the model. However, the substance is not within the structural domain (65%

unknown fragments). In addition the model issued a warning regarding the low water solubility. Therefore, the estimate is not reliable.

9.5 QPRF: Comparative analysis of estimated and measured BCF data (OECD 305; Müller & Nendza, 2011)

1.	Substance	CAS 103597-45-1		
2.	General information			
2.1	Date of QPRF	23 Sep. 2014		
2.2	QPRF author and contact details	BASF SE, Department for Product Safety, Ludwigshafen, Germany		
3.	Prediction			
3.1	Endpoint (OECD Principle 1)	Endpoint	Bioaccumulation (aquatic)	
		Dependent variable	Bioconcentration factor (BCF)	
3.2	Algorithm (OECD Principle 2)	Model or submodel name	Comparative analysis of estimated and measured BCF data (OECD 305)	
		Model version	Müller & Nendza (2011)	
		Reference to QMRF		
		Predicted value (model result)	see Table 14	
		Input for prediction	Log Kow	
		Descriptor values	Log Kow	
3.3	Applicability domain (OECD principle 3)	Domains:		
		Model	Range of log Kow	Within range
		1) Veith et al. (1979)	1 - 7.05; recommended range: 0 - 6	No (not within recommended range)
		2) Connell and Hawker (1988)	2.6 - 9.8	No
		3) European Communities (2003)	2.6 - 9.8; recommended range: 6 - 9.8	No (not within recommended range)
		4) Nendza (1991)	1 - 11	No
		5) Mackay (1982)	1 - 7.1	No
		6) Veith and Kosian (1983)	1 - 6.9	No
		7) Bintein et al. (1993)	1.2 - 8.5; recommended range: 6 - 8.5	No (not within recommended range)
		8) Schüürmann and Klein (1988)	1.8 - 6.5	No
		9) Könemann and van Leeuwen (1980)	3.5 - 6.4	No; Substance is not a chlorobenzene.

		10) Lu et al. (1999)	1 - 7.1	No (based on log Kow); although substance is a non-polar compound.
		11) Escuder-Gilabert et al. (2001)	0.3 - 5.8	No
		12) Neely et al. (1974)	2.6 - 7.6	No; Substance is not a halogenated aromatic.
		13) Zok et al. (1991)	0.9 - 2.8	No; Substance is not a substituted aniline.
3.4	The uncertainty of the prediction (OECD principle 4)	Model no. 1: heterogeneous dataset (Pimephales promelas); n = 55; r = 0.95 Model no. 2: heterogeneous dataset (fish (various)); n = 45 Model no. 3: heterogeneous dataset (fish (various)); n = 43; r = 0.883 Model no. 4: heterogeneous dataset (fish (various)); n = 132; model not derived by regression; therefore no statistical data available Model no. 5: heterogeneous dataset, mainly chlorinated hydrocarbons (fish (various)); n = 44; r = 0.95; s = 0.25 Model no. 6: heterogeneous dataset, mainly halogenated compounds (fish (various)); n = 122; r = 0.927; s = 0.49 Model no. 7: heterogeneous dataset (fish (various)); n = 154; r = 0.95; s = 0.347 Model no. 8: heterogeneous dataset, mainly chlorinated and polycyclic hydrocarbons (fish (various)); n = 32; r = 0.87; s = 0.54 Model no. 9: chlorobenzenes (*Poecilia reticulata*); n = 6; r = 0.999; s = 0.039 Model no. 10: diverse non-polar chemicals (various fish); n = 80; r = 0.944 Model no. 11: diverse (various fish); n = 66; r = 0.917 Model no. 12: halogenated aromatics (*Salmo gairdneri*); n = 8; r = 0.949 Model no. 13: substituted anilines (*Brachydanio rerio*); n = 9; r = 0.934		
3.5	The chemical mechanisms according to the model underpinning the predicted result (OECD principle 5)	Quantitative structure-activity relationships (QSAR) make use of the fact that bioaccumulation of stable organic compounds is governed by partitioning between aqueous and lipid phases. The predominant process of passive diffusion is frequently formalized in log Kow-dependent QSAR models. It is often assumed, that the log Kow-based BCF estimates represent a 'worst case' reference point. Estimating bioconcentration factors (BCF) from octanol/water partition coefficients (log Kow) is well established and essentially valid for neutral organics of intermediate		

		lipophilicity (0 < log KOW < 6) (European Communities, 2003; Nendza, 1991; Nendza, 1998; Dearden, 2004).

References

- Bintein S, Devillers J, Karcher W. 1993. Nonlinear Dependence of Fish Bioconcentration on n-Octanol/Water Partition Coefficient. SAR QSAR Environ. Res. 1: 29-39.
- Connell DW, Hawker DW. 1988. Use of Polynomial Expressions to describe the Bioconcentration of Hydrophobic Chemicals by Fish. Ecotox. Environ. Saf. 16: 242-257.
- Dearden JC. 2004. QSAR modeling of bioaccumulation. In Predicting chemical toxicity and fate, ed. Cronin, M. T. D. and Livingstone, D. J. 333-55. Boca Raton: CRC Press.
- Escuder-Gilabert L, Martin-Biosca Y, Sagrado S, Villanueva-Camanas RM, Medina-Hernandez MJ. 2001. Biopartitioning Micellar Chromatography to Predict Ecotoxicity. Analytica Chimica Acta 448: 173-185.
- European Communities. 2003. Technical guidance document on risk assessment in support of Commission Directive 93/67/EEC on risk assessment for new notified substances, Commission Regulation (EC) No 1488/94 on risk assessment for existing substances, Directive 98/8/EC of the European Parliament and of the Council concerning the placing of biocidal products on the market. Joint Research Centre, Ispra, Italy: European Commission.
- Köneman H, van Leeuwen K. 1980.Toxicokinetics in Fish: Accumulation and Elimination of Six Chlorobenzenes by Guppies. Chemosphere 9: 3-19.
- Lu XX, Tao S, Cao J, Dawson RW. 1999. Prediction of Fish Bioconcentration Factors of Nonpolar Organic Pollutants based on Connectivity Indices. Chemosphere 39: 987-999.
- Mackay D. 1982. Correlation of Bioconcentration Factors. Environ. Sci. Technol. 16: 274-278.
- Müller M, Nendza M (2011). Comparative analysis of estimated and measured BCF data (OECD 305). Federal Environment Agency (Umweltbundesamt), Texte 15/2011, Report no. UBA-FB 001435/E . 54 pp.
- Neely WB, Branson DR, Blau GE. 1974. Partition Coefficients to Measure Bioconcentration Potential of Organic Chemicals in Fish. Env. Sci. Technol. 8: 1113-1115.
- Nendza M. 1991. QSARs of bioconcentration: validity assessment of log Pow/log BCF correlations. In Bioaccumulation in aquatic systems, ed. Nagel, R. and Loskill, R. 43-66. Weinheim: VCH.
- Nendza M. 1998. Structure-activity relationships in environmental sciences. London, Great Britain: Chapman & Hall.
- Schüürmann G, Klein W. 1988. Advances in Bioconcentration Prediction. Chemosphere 17: 1551-1574.
- Veith GD, Defoe DL, and Bergstedt BV. 1979. Measuring and estimating the bioconcentration factor of chemicals in fish. J.Fish.Board Can. 36: 1040-1048.
- Veith GD, Kosian P. 1983. Estimating Bioconcentration Potential from Octanol/Water Partition Coefficients. In: Physical Behaviour of PCBs in the Great Lakes. Mackay D, Paterson S, Eisenreich SJ, Simmons MS (Eds.), Ann Arbor Science Publishers, Ann Arbor, MI, U.S.A.
- Zok S, Görge G, Kalsch W, Nagel R. 1991.Bioconcentration, Metabolism, and Toxicity of Substituted Anilines in the Zebrafish (*Brachydanio rerio*). Sci.Tot. Environ. 109/110: 411-421.