codex alimentarius commission

FOOD AND AGRICULTURE ORGANIZATION OF THE UNITED NATIONS

WORLD HEALTH ORGANIZATION

JOINT OFFICE: Viale delle Terme di Caracalla 00100 ROME Tel: 39 06 57051 www.codexalimentarius.net Email: codex@fao.org Facsimile: 39 06 5705 4593

Agenda Item 7 **CX/MAS 05/26/9-Add.1**

JOINT FAO/WHO FOOD STANDARDS PROGRAMME

CODEX COMMITTEE ON METHODS OF ANALYSIS AND SAMPLING
Twenty-sixth Session
Budapest, Hungary, 4-8 April 2005

CONSIDERATION OF THE METHODS FOR THE DETECTION AND IDENTIFICATION OF FOODS DERIVED FROM BIOTECHNOLOGY

GENERAL APPROACH AND CRITERIA FOR THE METHODS

COMMENTS (United States, AOCS)

UNITED STATES

The United States appreciates the complexities of the document that Germany and the United Kingdom have drafted. This document identifies basic criteria for evaluating DNA-based and protein-based methods for identifying the presence of foods derived from modern biotechnology.

The document identifies many of the critical issues associated with both DNA-based and protein-based methods. The United States has both general and specific comments with respect to this document. General comments are listed below; specific comments relative to the document are reflected in the attached Word document file with the recommended modifications highlighted (tracked changes).

General Comments

1. This document does not adequately reflect the importance of protein-based testing in the world. Protein-based tests are generally less costly, simple to use, and amenable to use in non-laboratory environments. The United States has added language to the document (attached) that we feel addresses this issue.

2. The terms “GMO”, “genetically modified”, and “GM” are used throughout this document. In other Codex committees, these terms have been replaced with terms such as “foods derived from modern biotechnology” and “recombinant DNA.” To be consistent within Codex, the United States strongly recommends that “GMO”, “genetically modified”, and “GM” be eliminated entirely from this document. The United States has added language to the document (attached) that we feel addresses this issue.

3. Historically, CCMAS has not endorsed methods for which there are no Codex standards or provisions. Since standards or provisions for foods derived from modern biotechnology have not been established by the responsible Codex Committees, it would be inappropriate for CCMAS to endorse any methods to detect foods derived from modern biotechnology. CCMAS made it clear in its 2002 Meeting Report (Alinorm 03/23, paragraph 86) “that the selection or endorsement of methods without appropriate provisions was not possible” and this should be made clear in the background discussion for this document.

4. The United States does not believe that this document should address how a particular methodology can be used nor should it discuss policy issues associated with labeling or the approval processes.

AOCS

Comments on CCMAS document CX/MAS 05/26/9 from AOCS

This document, developed under the leadership of the UK and German delegation, provides an important and impressive guidance for criteria and validation of DNA-based detection tools. We are grateful to be given the opportunity to forward the comments below.

The following general comments have been developed:

1. Protein-based methods require greater acknowledgement, representation and development within the document.
2. The terms GMO, genetically modified, and product of modern biotechnology are used interchangeably in the document. The authors might consider use of a single term, or specifying that the three are interchangeable if they consider them so.

Specific technical comments are as follows: inserts are shown as {_}, deletions as [ ]

APPENDIX 1:

Paragraph 4: However, the two most common approaches are those based on DNA-based {detection} methods and those based on the detection of protein. The former is generally performed via PCR, although other methods that achieve measurement without a PCR step may be employed if properly validated. Both DNA and protein-based approaches are considered here, [though it is the DNA-based PCR approach which is generally recognised as being the more widely applicable.]

DELETE: "though it is the DNA-based PCR approach which is generally recognised as being the more widely applicable"
REASON: Both DNA and Protein methods are applicable and the method to choose depends on the situation. It is almost certainly the case that more protein-based tests are carried out than DNA based tests, and therefore it could be argued that "protein methods are more widely applicable".

METHOD DEVELOPMENT TO……
Applicability of the Method, 2nd paragraph:
As an example it is required from {a DNA}extraction method, independent of matrix to which it is to be applied, that it yields DNA of sufficient quantity, structural integrity and purity to allow a proper evaluation of the performance of the subsequent method steps (e.g. adequate amplification of DNA during the PCR step) to be undertaken. [T]{For a real-ime PCR method, t}his can be tested, for example, by setting up dilution series of the template DNA and determining that the ΔCT in a real-time PCR analysis between the dilutions corresponds to the dilution factor[, e.g.]{. For example,} if DNA is diluted 10X then the ΔCT should be approx. 3.32, if the DNA is diluted 4X, the ΔCT should be 2, etc. Deviations from this relationship may indicate that the extracted DNA contains PCR inhibitors, that the DNA solution is not homogenous or the DNA quantity so low that stochastic variation in copy numbers yield unreliable quantitative estimates.

Add the following paragraph concerning Protein methods
{For a protein method, extraction procedures should be expected to yield extracts of sufficient stability that the protein concentration will not change significantly during the analysis. As with the DNA methods, the presence of inhibitors or interfering agents can be assessed by analysing serial dilutions of the extract when using quantitative methods. In addition, some estimation of the total protein extracted, and efficiency of extraction may be desirable.}

Section: **Modular Approach to Method Validation**

Paragraph 1: The "method" refers to all the experimental procedures needed to estimate the measurand in a particular matrix. For {DNA methods with} a particular material, this may include the methods for DNA extraction and the final quantification in a PCR system. In such a case, the whole chain from extraction up

to the PCR-method (or equivalent) constitutes a method, but the different method parts can be considered separately (i.e. modular validation). In practice, this is difficult to achieve.

Add at end of section: {The Modular Validation approach is not appropriate to protein based methods, and has never been applied to such methods.}

Section: **General Information**

The provision of the detection method is aimed to serve mainly the requirements for the monitoring and labelling of GMOs, as set out in the specific regulations. To serve these purposes, the method should detect and quantify the specific GM event in the GM product; this may be achieved using either protein-based or DNA-based methods.

Currently, the DNA-based detection method typically consists of PCR methodology and includes:

Change 5th bullet to

- appropriate {reference and/or}control samples.

Sub-Paragraph (1) GMO Screening Methods. In the case of a method used for screening for the presence of multiple events, the method should be specific and allow for unequivocal detection/identification/ quantification of a specific DNA sequence in the case of DNA-based methods. [In the case of Protein based methods, the method should be specific and allow for unequivocal detection/identification/ quantification of a specific protein.] – delete last sentence and move to end of section

Add paragraph to end of section:
{Currently, the protein-based detection method typically consists of ELISA or other antibody-based methodology in a multi-well plate, glass tubes, or a lateral flow strip and includes:

- a protocol describing an extraction method which is applicable to a relevant matrix;
- a source of the antibodies required (usually supplied as a commercial kit);
- a protocol describing the conditions under which the analysis can be used to detect the target protein;
- appropriate reference and/or control samples.

In the case of Protein based methods, the method should be specific and allow for unequivocal detection/identification/ quantification of a specific protein. However, where a protein or specific combination of proteins is characteristic of a particular event, the protein method may be used to identify material containing that event.}

Section: **Collaborative Trial Test Materials**

Paragraph 2: In other fora recommendations have been made that in case of “general purpose” GMO procedures.
It is not clear what this phrase “general purpose” refers to.

Header: **VALIDATION OF PCR METHODS**
Delete the term **PCR.** The header then reads: **VALIDATION OF METHODS**

Section: **UNITS OF MEASUREMENT**

Paragraph 3:
Change "only 0.29%" to "approximately 0.3%" (the value and methods used are not precise enough to use 0.29%)

Section: **REFERENCE MATERIALS**

Paragraph 2 and 3:
Codex may consider requiring the availability of suitable reference materials as part of the method endorsement procedure. However, it is recognised that there are specific problems with the development of reference materials {for some uses. For example, maize materials for use }[should the maize event or the] {with} construct specific methods [be considered,] {or for a method which recognizes an element present in many events (e.g. 35S.

A suitable reference material is generally required for validation of a method. Suitable reference materials are becoming available for many commercialized events. Where they are not available, the {best available material, such as} quality control materials from proficiency testing schemes or from the use of plasmid or amplicon DNA may be considered.

{For ELISA plate-based protein methods, the reference or control material is normally included in the kit. Lateral flow strips are typically validated extensively by the manufacturers, but the use of the best available control material is useful for confirmation of the result.}

REFERENCES

Add reference to GIPSA documents
Add reference to LIPP ET AL.: "Polymerase Chain Reaction Technology as Analytical Tool in Agricultural Biotechnology" JOURNAL OF AOAC INTERNATIONAL VOL. 88, NO. 1, 2005, pp136-155.

ANNEX 1:

Specification of the prediction model/mathematical model needed for the method

If the derivation of the results relies upon a mathematical relationship this must be outlined and recorded (e.g., a regression line or calibration curve obtained by other means). Instructions for the correct application of the model should be provided. These may include, depending on the method, a recommended number and range of levels to be analysed, minimum number of replicates to be included or the means to evaluate the goodness-of-fit[.] {as well as appropriate threshold values for the real-time PCR curves.}

INFORMATION ABOUT THE METHOD OPTIMISATION

Primer pairs tested

For PCR methods, sufficient justification should be given of how and why the proposed primer pair{s and (if used) probes have }[has] been selected[,] {this information should also be given }[also] for the reference gene (should this be part of the method).

Cross-reactivity

The cross-reactivity, interferences and matrix effects should be evaluated[, particularly]. {This applies to both DNA and} [for] the protein-based methods of analysis

(correct spelling of word 'interference' in next paragraph: *Applicability*)

METHOD VALIDATION/PERFORMANCE

- ***whether the method is instrument specific***

(Comment: it is prefereable if methods are able to be easily adapted to a number of different models and types of equipment so that the analyst is not limited to purchasing the required equipment from a single manufacturer)

At the moment a number of different types of real time instruments are available. These instruments may have different heating and cooling characteristics, which affect ramp rates and [affects] the time necessary for a whole PCR run.

Beside{s} the differences in the heating and cooling system{s}, there are differences in the technique used to induce and subsequently to record the fluorescence. Some real time instruments use laser technique for inducing fluorescence, others are equipped only with a white lamp and filters for selecting a specific wave length. The detection of the fluorescence could also vary.

Taking all the differences into account it is[impossible]{sometimes difficult} to change the instrument without adaptation of the PCR method. Thus, {the method developer should attempt to develop methods which are robust enough to performed on a range of instrument types}. Methods that are very instrument dependant are less desirable than those which are generally applicable [can be because the methods are generally] although [instrument dependent they] methods cannot be transferred to other equipment without evaluation and/or modification {and appropriate in-house validation}.

[This is in many ways equivalent to the Codex Type I method and should be considered in the same light.] – delete as there are, or will be reference materials available which allow comparison of results with real samples of known composition.

Delete the following as this annex deals only with PCR.
This material may be more appropriate in the body text of the Appendix 1.

[And for both protein and DNA based methods:

- ***whether there are differences between PCR-based and immunological methods concerning validation criteria***

The DNA and protein-based techniques used to detect and quantify a GMO derived material in foods are based on different principles. In PCR the targeted DNA is amplified in a exponential manner, in which a small difference in the beginning of the PCR process will lead to a big difference in the amplified amount of DNA after 35-45 cycles. In contrast [to that] immunological detection assays are based on the direct interaction with the target molecule and do not include an amplification step.

Moreover, the quantitation by real time PCR is often based on two independent PCR systems: one for the genetic modification and one for the taxon specific sequence {and the uncertainty/method characteristics of each should be considered}.]

ANNEX III: VALIDATION OF A QUANTITATIVE PCR METHOD

Section: **Dynamic Range - Range Of Quantification**

Recommendation: The dynamic range of the method should cover at least [20% and 5] times the target concentration, where practicable. Target concentration should be understood here as the threshold relevant for a certain regulation.
The range the method should cover is stated in various terms in different sections of the document – either as 20% and 5 or 10% and 5 times the target concentration. As the target concentration may vary according to local regulations, the range should be defined by the method itself, and not in terms of a 'target ' which is not globally relevant. Is this generally the case that that a range either side of the 'target' is desired to be shown effective for a particular method? This issue arises in many places in the document such as in LOD and LOQ.

Section: Limit of Detection (LOD)
Paragraph 2: last sentence: However, the LOQ shall always be established and included in the validation study {for a quantitative method}..

Paragraph 3:5th sentence For [quantitative] {qualitative} methods the LOD is the amount of analyte at which the analytical method detects the presence of the analyte at least 95% of the time (≤5% false negative results).

RSDr and RSDR: A reasonable value for these is yet to be determined, and it is considered that it is not possible to define these without a significant amount of empirical data from actual global validation studies.

Ruggedness (Robustness):

It is not clear how recovery can be expressed in this type of PCR method.

Sensitivity:

Dilution of a 5% CRM may not be suitable for cases where the regulated level is at or above 5%.
We suggest the wording {Alternatively a suitable CRM may be diluted without target DNA}
This sentence might be better placed at the end of the preceding paragraph.

Specificity:

Recommendation: Specificity is the starting point for a method and needs to be considered during primer {and probe} design. Primers {and probes} should be checked against the known sequence of the event insert and pertinent databases for possible matches. Specificity must also be demonstrated experimentally. The following suggests a reasonable approach and the experiments should be performed during pre-validation of an assay.

It might be useful to define 'significant' instrument reading, as with these methods, a spurious reading might be observed if enough PCR cycles are performed.

CONSIDERATION OF THE METHODS FOR THE DETECTION AND IDENTIFICATION OF FOODS DERIVED FROM BIOTECHNOLOGY

GENERAL APPROACH AND CRITERIA FOR THE METHODS

BACKGROUND

At the Twenty-fourth Session of the Codex Committee on Methods of Analysis and Sampling, papers giving the methods that had been collated by the *ad hoc* Intergovernmental Task Force on Food Derived from Biotechnology (see CX/MAS 02/8) and outlining general considerations of methods of analysis for the detection and identification of foods derived from biotechnology (see CX/MAS 02/9) were discussed. It was noted that the presence of ~~genetically modified~~**recombinant DNA** organisms or their derivatives could be assessed by the detection of either DNA sequences present as a result of recombination or the protein coded by the inserted gene. It was pointed out that protein-based methods were cheap, offered high selectivity and sensitivity but that since proteins were denatured during processing these techniques were most suitable for the analysis of raw materials and were not generally applicable to highly processed foods. It was also noted that these methods cannot be used when no new protein is expressed in the food, and these methods cannot differentiate between genetic events that produce the same protein.

Methods of detection of DNA markers based on the polymerase chain reaction (PCR) have been used in a variety of food analyses and widely used for detection of ~~GM~~ **recombinant DNA** derivatives in food for many years, and modifications of the PCR method were also widely used. A typical method involved several steps such as sampling, extraction and purification, amplification by PCR and detection/quantification. Specific questions arising in the area of proficiency testing, use of performance criteria and the necessity of quantification due to threshold settings since the results of investigations showed the difficulties in measuring low levels of ~~GM~~ **recombinant DNA** material in processed foods were also discussed. Methods described in the collated documents could only be used successfully if all information about the sequence and certified reference materials were available.

GENERAL CRITERIA

In view of the absence of precise provisions for ~~GMOs~~ **foods derived from modern biotechnology** and of difficulties with the practical application of methodology in this area, the Committee proposed to develop recommendations with respect to criteria for methods of analysis and for quality control measures that should be introduced in laboratories offering ~~GM~~ analyses **for foods derived from modern biotechnology**. It was agreed that a Working Group led by Germany and the United Kingdom would update and further develop the paper for this session and prepare recommendations for quality control measures in laboratories and criteria for methods of analysis for the Twenty-fifth Session of CCMAS.

The paper CX/MAS 04/10 was discussed at the Twenty-fifth Session of CCMAS, where the following comments were made or were noted:

- The Committee recalled that the last session had agreed that the Delegations of Germany and the United Kingdom in cooperation with a drafting group would prepare a revised document that would include recommendations for quality control measures in laboratories and criteria for methods of analysis.
- The Delegation of the United Kingdom introduced the document and indicated that it included recommendations on the criteria for methods of analysis and quality control measures that should be introduced in laboratories performing ~~GM~~ analysis **for foods derived from modern biotechnology**, with specific focus on the detection of DNA markers based on PCR that were more commonly used.
- The Delegation of Germany referred to the list of methods developed by the Task Force on Foods Derived from Biotechnology and highlighted the importance of further work on guidelines that would provide guidance to governments to select methods for the detection of foods derived from biotechnology.
- The Delegation of the United States welcomed the paper that provided a good scientific basis for further discussion and drew the attention of the Committee to its comments in CRD 9. It noted in particular that the document developed criteria mostly for DNA-based methods but that alternative methods based on the detection of protein should also be addressed.

- The Delegation of Brazil expressed the view that the validation of immunoassay methods should be considered, and that in Annex 1 more information should be included on the description of the method, such as: complete description of the primer, number of cycles, composition of cycles, equipment, amplicon length, type of polymerase and reference material.
- The Delegation of Japan questioned the application of those criteria contained in the document to the detection of ~~GMOs~~ **foods derived from modern biotechnology** although they are applicable to chemical analysis.
- The Delegation of Norway proposed to amend the section on the modular approach to reflect that it should not be used "unless independence between the modules can be documented", since it should not be systematically avoided.
- The Delegation of Cuba drew the attention of the Committee to the issues related to consumer protection, that might need to be addressed by the Task Force in the future and in particular the level of transgenicity of the material.
- The Committee discussed whether new work should be initiated in the Step Procedure in order to circulate for comments as soon as possible the document in Appendix I: *Guidelines for the Validation and Quality Control Requirements for ~~GMO~~ Analyses* ***for Foods Derived from Modern Biotechnology.***
- Some delegations stressed the need to proceed rapidly as governments needed guidance on this very important and complex issue. Other delegations indicated that they had been part of the original Working Group but there had not been enough time to provide detailed comments and that it would be preferable to consider the text carefully before initiating the elaboration of specific guidelines.
- The Committee agreed that the document would be revised by the Delegations of the United Kingdom and Germany with the assistance of a Drafting Group for consideration at the next session, with a view to the elaboration of Guidelines.

The following countries and organisations expressed their willingness to participate in this work: Argentina, Australia, Brazil, Canada, Egypt, France, Iran, Ireland, Italy, Japan, Malaysia, The Netherlands, Norway, Philippines, United States, European Commission, AOAC International, AOCS, Bio, CROPLIFE International, EUROPABIO, and ISO.

These measures are given as Guidelines in the Appendix to this paper.

RECOMMENDATIONS

It is recommended that the draft Guidelines be discussed at the Twenty-sixth Session of CCMAS. If there is sufficient consensus, then the approaches described should be further refined and then sent to governments for comment and progress through the Codex system.

APPENDIX I: GUIDELINES FOR THE VALIDATION AND QUALITY CONTROL REQUIREMENTS FOR THE ANALYSIS OF FOODS DERIVED FROM MODERN BIOTECHNOLOGY

INTRODUCTION

Method Criteria

The Codex Alimentarius Commission places an emphasis on the acceptance of methods of analysis which have been "fully validated" through a collaborative trial conforming to an internationally accepted protocol. In a number of sectors, including the foods derived from **modern** biotechnology ~~(GMO)~~ sector, there are few methods of analysis which have been fully validated. As a result, Codex is also endorsing by reference single-laboratory validation protocols. In this area there may be pressure to adopt a formal single-laboratory validation as an interim measure in the absence of collaborative trial data. However, methods used for determination of the presence of ~~GMO's~~ **foods derived from modern biotechnology** are able to be, and intended to be performed at, multiple laboratories and should therefore be validated by multi-laboratory collaborative studies as soon as practicable.

~~In these Guidelines the term "GMO" has been used for "Foods Derived from Biotechnology".~~

Many methods are currently being developed for ~~GMO~~ detection **of foods derived from modern biotechnology**, identification and quantification. Before they are accepted for use by Codex they must be validated to ensure that they are fit-for-purpose.

However, the two most common approaches are those based on DNA-based **detection** methods and those based on the detection of protein. The former is generally performed via PCR, although other methods that achieve measurement without a PCR step may be employed if properly validated. Both DNA and protein-based approaches are considered here~~, though it is the DNA based PCR approach which is generally recognised as being the more widely applicable~~.

The conventional criteria that have been adopted by Codex for the evaluation of methods of analysis are:-

- accuracy
- applicability (matrix, concentration range and preference given to 'general' methods)
- limit of detection
- limit of determination
- precision; repeatability intra-laboratory (within laboratory), reproducibility inter-laboratory (within laboratory and between laboratories)
- recovery
- selectivity
- sensitivity
- linearity

These Guidelines address these requirements ~~in the GMO sector~~, and anticipates that **it** is likely that these will have to be further expanded (e.g. for PCR) by other items such as:-

- amplicon length
- whether the method is instrument specific
- whether there are differences between qualitative and quantitative PCR-based detection methods
- whether single- or multi-plex PCR amplifications are undertaken

for the DNA-based methods.
And

- equivalency of reagents over time

for the protein based methods

The method validation process accepted by Codex includes the definition of the requirements for the method, testing that the method meets these requirements when carried out, for instance, by different laboratories in different countries, and documentation of the method performance and measurement uncertainty.

Criteria Approach

Codex Alimentarius Commission has accepted the "criteria approach" for methods of analysis. This approach does not extend to Codex Type I empirical/defining, procedures. It is necessary to ensure that this approach is incorporated into Codex guidelines on the validation of ~~GMO~~ methods of analysis **for foods derived from modern biotechnology** unless it is explicitly stated that all ~~GMO~~ **such** methods of analysis are empirical, both theoretically as well as in practice.

Laboratory Quality

The Codex Alimentarius Commission has adopted guidelines for the "quality" of laboratories involved in the import and export of foods. These quality characteristics are based on accreditation to ISO/IEC Standard 17025, proficiency testing and internal quality control as well as the use of methods of analysis validated according to Codex requirements. These overarching guidelines provide information to and dictate requirements for laboratories **analyzing for foods derived from modern biotechnology**~~working in the GMO sector~~.

Measurement Uncertainty

Codex is currently developing guidelines on Measurement Uncertainty. These guidelines, as well as the accreditation requirements cited above, require laboratories to estimate the uncertainty of their quantitative measurements. This is particularly important and has consequences for measurements ~~in the GMO sector~~ where analytical controls may not be as effective as found in other areas of analysis in the food sector. It is frequently not appreciated that the magnitude of the measurement uncertainty is considerably greater in this analytical sector than would normally be expected.

INFORMATION TO BE PROVIDED TO CODEX WHEN A METHOD FOR ~~GMOs~~ THE DETECTION OF FOODS DERIVED FROM MODERN BIOTECHNOLOGY IS TO BE CONSIDERED FOR ENDORSEMENT BY CCMAS

The information that should be supplied to CCMAS when a method is to be considered for endorsement is given in Annex I. The annex lists both general considerations and specific requirements.

As ~~GMO~~ methodology **for foods derived from modern biotechnology** becomes more developed the specific requirements will be converted to performance criteria to conform to the "criteria approach" already adopted by Codex.

DEFINITIONS

There are a number of Codex definitions applicable to ~~GMO~~ analysis **for foods derived from modern biotechnology**. Suggested definitions are given in Annex II.

METHOD DEVELOPMENT TO FORMAL VALIDATION

Applicability of the Method

This is a particularly important criterion in ~~GMO~~ analysis **for foods derived from modern biotechnology**. In principle the method should be applicable to the matrix of concern within the Codex system. If this is a specific product ~~derived from GMO~~ then there is merit in requiring those seeking endorsement to provide information on the method of analysis appropriate to the specific product and, ideally, the matrix in which it is likely to be used. In case of "general purpose" ~~GMO~~ methods **for foods derived from modern biotechnology**, at least one extraction method applicable to a general matrix should be available.

As an example it is required from a~~n~~ **DNA** extraction method, independent of matrix to which it is to be applied, that it yields DNA of sufficient quantity, structural integrity and purity to allow a proper evaluation of the performance of the subsequent method steps (e.g. adequate amplification of DNA during the PCR step) to be undertaken. ~~This~~ **For a real time PCR method, this** can be tested, for example, by setting up dilution series of the template DNA and determining that the ΔCT in a real-time PCR analysis between the dilutions corresponds to the dilution factor, e.g. if DNA is diluted 10X then the ΔCT should be approx. 3.32, if the DNA is diluted 4X, the ΔCT should be 2, etc. Deviations from this relationship may indicate that the extracted DNA contains PCR inhibitors, that the DNA solution is not homogenous or the DNA quantity so low that stochastic variation in copy numbers yield unreliable quantitative estimates.

For a protein method, extraction procedures should be expected to yield extracts of sufficient stability that the protein concentration will not change significantly during the analysis. As with the DNA methods, the presence of inhibitors or interfering agents can be assessed by analyzing serial dilutions

of the extract when using quantitative methods. In addition, some estimation of the total protein extract, and efficiency of the extraction is desirable.

Validation Process

Method validation is a process of establishing the performance characteristics and limitations of an analytical method and the identification of the influences, which may change these characteristics - and to what extent. The results of a validation process describe which analytes can be determined in what kind of matrices in the presence of which interference. The validation exercise results in precision and accuracy values of a certain analytical method under the examined conditions.

Formal validation of a method is the conclusion of a long process, which includes the following main steps:

- ***Method development and optimisation***. Prior to any pre-validation, the method should be fully optimised so that an inter-laboratory transfer is possible. The protocol should be finalized so that no major changes are needed between the pre-validation and validation.

- ***Pre-validation of the method***. Pre-validation should ensure that a method performs in a manner, which allows a successful conclusion of the validation study, i.e. it should provide evidence about the compliance with the regulations. Pre-validation should preferably be carried out by involving 2 – 4 laboratories.
- ***Full validation of the method.*** Full validation requires considerable resources and should be conducted only on methods which have received adequate prior testing.

A collaborative trial is expensive to undertake and usually follows only after the method has shown acceptable performance both in a single-laboratory and a pre-validation study.

Modular Approach to Method Validation

The "method" refers to all the experimental procedures needed to estimate the measurand in a particular matrix. For **DNA methods with** a particular material this may include the methods for DNA extraction and the final quantification in a PCR system. In such a case, the whole chain from extraction up to the PCR-method (or equivalent) constitutes a method, but the different method parts can be considered separately (i.e. modular validation). In practice this is difficult to achieve.

The theoretical advantage of a modular approach to method validation is that each section of a method or protocol can be validated separately, and once validated, can be combined with other sections in a flexible manner.

However, there are several disadvantages to a modular approach to method validation, particularly when ~~GMO~~ analysis **for foods derived from modern biotechnology** is being considered. It has been found that the variability of ~~GMO~~ **this** analysis is very significant, and this then reduces the effectiveness in comparing different approaches to the same module in a method. But most critically, a modular approach to method validation assumes that the modules in a method which form the whole are independent of each other. This is frequently not the case and where "official control" work is to be undertaken, a modular approach should not be taken unless independence between the modules can be clearly demonstrated and documented.

The modular approach is not appropriate to protein-based methods.

Note: The Modular Approach to Method Validation is very suspect. As described in this section, there are several significant concerns, particularly with respect to DNA-based methods to detect foods derived from modern biotechnology. Given the uncertainties with the Modular Approach, we suggest this section be deleted from the document.

METHOD ACCEPTANCE CRITERIA

In order to evaluate a method prior to full validation, information concerning both the method and the method testing is required. Information on this is given in Annex I.

The method will be evaluated based on the information provided to Codex. The evaluation should verify that the princip**al**~~le~~ preconditions for using the method for Codex purposes are fulfilled. This section describes the method acceptance criteria, which have to be fulfilled by the method in order to conduct further a pre-validation and full collaborative trial.

~~**Principle**~~ **Principal Conditions**

~~The provision of the detection method is aimed to serve mainly the requirements for the monitoring and labelling of GMOs, as set out in the specific regulations. To serve these purposes, the~~ **The purpose of the** method **is to** ~~should~~ detect and quantify the ~~specific GM~~ **recombinant DNA** event in the ~~GM~~ product~~;~~ **and** this may be achieved using either protein-based or DNA-based methods.

Currently, the DNA-based detection method typically consists of PCR methodology and includes:

- a protocol describing an extraction method which is applicable to a relevant matrix;
- a description of the oligonucleotide primer sequences which uniquely identify the ~~GM~~ **recombinant DNA** event in the ~~GM~~ product[1];
- a description of the oligonucleotide primer sequences which amplify an endogenous gene sequence applicable to the specific host species;
- a protocol describing the conditions under which PCR can be used to detect the **recombinant DNA event**~~GM product~~;
- appropriate **reference and/or** control samples.

The method provider should demonstrate that the method fulfils the ~~principle~~ **principal** method requirements:

(1) ~~GMO~~ Screening Methods **for foods derived from modern biotechnology**. In the case of a method used for screening for the presence of multiple events, the method should be specific and allow for unequivocal detection/identification/ quantification of a specific DNA sequence in the case of DNA-based methods. ~~In the case of Protein-based methods, the method should be specific and allow for unequivocal detection/identification/ quantification of a specific protein.~~

(2) DNA-based event-specific methods should allow for unequivocal detection/identification/ quantification of a known target nucleotide sequence.

Currently, the best choice concerning event-specificity of a method, should PCR be the chosen technique, is to target an event-specific genomic region using a set of oligonucleotides (primers) that trigger the amplification of such a region. Among various types of event-specific genomic regions, the one relative to the junction between the **recombinant DNA**~~transgenic~~ insert and the host genomic DNA will probably be the location of choice. However, when a unique DNA sequence can be found within the **recombinant DNA**~~transgenic~~ insert, such a sequence can also be targeted by appropriate oligonucleotide primers and amplified through a PCR.

(3) **In the case of protein-based methods, the method should be specific and allow for unequivocal detection/identification/ quantification of a specific protein. Currently, the protein-based detection method typically consists of Enzyme-Linked Immunosorbent Assay (ELISA) or other antibody-based methodology in a multi-well plate, glass tubes, or a lateral flow strip format, and includes:**

- **A protocol describing an extraction method which is applicable to a relevant matrix;**
- **A source of the antibodies required (usually supplied as a commercial kit);**
- **A protocol describing the conditions under which the analysis can be used to detect the target protein; and**
- **Appropriate reference and/or control samples.**

In the case of protein-based methods, the method should be specific and allow for unequivocal detection/identification/quantification of a specific protein. However, where a protein or specific combination of proteins is characteristic of a particular event, the protein method may be used to identify material containing that event.

(4) All methods should be applicable to the material specified in their scopes, and to appropriate quality control and reference materials.

[1] *Note: the fact that most event-specific sequences are not publicly disclosed should be discussed by CCMAS.*

It should be noted that at present only relative quantitation can be carried out, which means that the **recombinant DNA**~~transgenic~~ material relative to the corresponding ingredient/species is measured.

COLLABORATIVE TRIAL REQUIREMENTS

General Information

The purpose of a collaborative trial is to fully validate the data provided by previous testing in a pre-validation or a single laboratory exercise and to determine methodological precision in terms of repeatability and reproducibility.

The values of any performance parameters reported from validation studies must be interpreted and compared with care. The exact values and their interpretation may depend – besides the performance of the method - on the extent of the method (e.g. a real-time quantitative PCR only versus a method chain ranging from extraction to the real-time PCR quantification), experimental design applied, exact calculation forms used to determine the parameters and the approach used to detect and analyse outliers. In order to have meaningful "minimum performance requirements" the above factors must be treated appropriately and in a standardized manner.

For Codex purposes the ISO/AOAC/IUPAC harmonized protocol[1] has been adopted.

Minimum Performance Requirements

In a collaborative trial, the method performance should comply with the relevant parts of the method acceptance criteria and fulfil the method performance requirements specifically set below for the collaborative trial. Thus, the collaborative trial confirms the results obtained during the previous method evaluation phases and provides additional information about the method performance in a multi-laboratory setting. In particular, the compliance with the criteria for sensitivity and repeatability standard deviation should be re-confirmed.

In addition to the method acceptance criteria, at least the method performance requirements listed in Annex I should be evaluated from the experimental data of a collaborative trial. First, the definition and thereafter the requirements are described.

The endorsed methods and their associated validation data will be revised on a regular basis as the scientific knowledge and experience gained in Single-Laboratory validation and collaborative trials evolve. These Guidelines will also be complemented with practical information about the operational steps of the validation process.

Collaborative Trial Test Materials

In principle, the method should be applicable to and tested on the matrix of concern (i.e. on which any specification has been made).

In other fora recommendations have been made that in case of "general purpose" ~~GMO~~ procedures (in contrast to consideration of a specific product ~~derived from GMO~~) that the validation of the detection module is carried out using genomic DNA as the analyte (for a PCR-based method). This allows the detection step to be combined with various extraction methods applicable to different matrices. However, real materials/matrix typical of a type/group of matrices are preferred unless the effects of the materials/matrix on DNA quality in the extraction step is completely evaluated prior to applying a modular approach. Otherwise a modular approach is inappropriate when considering Codex specifications.

Note: The meaning of "general purpose procedures" needs to be clarified.

VALIDATION OF ~~PCR~~ METHODS

Specific information on the validation of quantitative and qualitative PCR methods is given in Annexes III and IV respectively.

Specific information on the validation of quantitative, semi-quantitative and qualitative protein-based methods is given in Annex V.

UNITS OF MEASUREMENT

~~Various countries have thresholds established for labelling of food and feed derived from modern biotechnology. These thresholds are explicitly or implicitly expressed as weight by relative percentage. However, n~~None of the current detection methods (DNA – or protein-based) are able to measure **the weight %**~~this~~ directly. Although there is a correlation between weight-% and the amount of DNA or protein,

respectively, the very nature of this relationship is influenced by a number of biological factors and thus remains highly variable. This continues to cause considerable misunderstanding and requires significant technical guidance.

Based on the PCR technique used for ~~GMO~~ **DNA target** identification and quantification genome equivalents are measured.

Therefore it is not trivial to consider how the ~~genetically modified~~ **recombinant DNA** material is calculated. For example, if a maize seed lot containing 2% ~~genetically modified~~ **recombinant DNA** seeds with the "new" trait in a hemizygous state (coming from the pollen) is used to prepare a flour sample then, in theory, ~~only 0.29%~~**approximately 0.3 %** of the isolated genomic DNA copies will represent the ~~genetically modified~~ **recombinant DNA** status. This is due to the different tissue types, the source from where the genomes in these tissue types are derived (maternal or paternal) and the contribution of the tissue types in the seed kernel. Consequently the amount of ~~genetically~~ **recombinant DNA** material would be underestimated (on a seed basis) by a DNA based approach to express the content of material derived from ~~genetically modified organisms~~ **recombinant DNA material**.

Note: The example given above in the 2nd sentence is based on assumptions that are not clearly stated. We would like to see a reference given for a full description of how the final value was obtained, or this sentence eliminated from the document.

Quantitation based on the "newly" expressed protein ~~in the GMO~~ would also lead to a significant contribution to the uncertainty of the analysis. For example the environment in which the material was grown can affect the amount of protein expressed. In addition, it is often the case that the protein is expressed at different levels in different tissue types of the plant. Consequently foods produced from different parts of a ~~genetically modified~~**recombinant DNA** plant would contribute a different amount of the "newly" expressed protein.

This issue needs to be appropriately addressed and performance and data reporting criteria established for these methods.

MEASUREMENT UNCERTAINTY

Analysts using methods which have been validated according to these guidelines will have available to them sufficient information to allow them to estimate the uncertainty of their result.

Guidance on the use of this measurement uncertainty estimation has been developed and adopted by Codex[2].

GUIDANCE ON LABORATORY SET-UP AND OPERATION

DNA-based methods for the analysis of foods derived from modern biotechnology apply techniques that are not considered as commonly available methods, as they require specific apparatus and handling techniques that differ from most chemical-analytical methods. It is therefore necessary to provide information and instructions on the essential differences in laboratory set-up and handling techniques. Examples are available[3].

REFERENCE MATERIALS

There are a number of matrices that can be used to develop reference materials or working standards for methods of detection of **recombinant DNA**~~GM~~ products. Each has its own advantages and drawbacks for particular purposes.

Codex may consider requiring the availability of suitable reference materials as part of the method endorsement procedure. However, it is recognised that there are specific problems with the development of reference materials **for some uses.** ~~, e.g. for~~ **For example,** maize materials **for use with construct specific methods, or for a method which recognizes an element present in many events [e.g. 35S, etc.)**~~should the maize event or the construct specific methods be considered.~~

A suitable reference material is generally required for validation of a method. Suitable reference materials are becoming available for many commercialized events. Where they are not available, the availability of quality control materials from proficiency testing schemes or from the use of Plasmid or amplicon DNA may be considered.

For ELISA plate–based protein methods, the reference or control material is typically included with the test. Lateral flow strips are typically validated extensively by the manufacturers, but the use of the best available control material is useful for confirmation of the result.

SAMPLING

In the area of ~~GMO~~ analysis **for foods derived from modern biotechnology,** it may be anticipated that sampling error can be expected to contribute significantly – if not dominate - the overall uncertainty of an analytical result, particularly when considering raw commodities. The combination of sampling and analytical uncertainties **has been addressed by CCMAS[1].** ~~is now being addressed by a number of International Organisations, most notably EURACHEM which has set up a new Working Group dealing with uncertainty of sampling. Much work has been carried out on sampling generally by CCMAS[4] and of bulk sampling for GMOs by the EU JRC[5], ISO/CEN[6] and GIPSA.~~

Note: We suggest that the "Codex General Guidelines on Sampling" be the only reference.

REFERENCES

1. ~~ISO/AOAC/IUPAC harmonized protocol (Protocol for the Design, Conduct and Interpretation of Method-Performance Studies, Ed. Horwitz, Pure & Appl. Chem. 331-343, **67**, 1995~~

~~2. Guidelines on the Use of Measurement Uncertainty Within Codex (being developed)~~

~~3. Draft ISO standard (ISO/DIS 24276) or the corresponding French standard (AFNOR XP V03-020-2, tabled as room document CRD 5 in its previous version AFNOR XP V03-020-1 by the French Delegation at the 24th Session of CCMAS)~~

14. Codex General Guidelines on Sampling.

5. ~~FP5 KeSTE project.~~

~~6. prEN ISO 21568~~

ANNEX I: INFORMATION TO BE PROVIDED TO CODEX WHEN A METHOD IS TO BE CONSIDERED FOR ENDORSEMENT BY CCMAS

In order to aid the endorsement of a proposed method of analysis ~~in the GMO sector~~**for foods derived from modern biotechnology** by Codex, and in particular CCMAS, the following should be provided:

DESCRIPTION OF THE METHOD

A complete and detailed description of all the components of the method should be provided. The use of multiple plates for PCR and protein methods, as an example, should be explicitly addressed. The information should also include information on the following:

Purpose and relevance of the method

The objective of the method and the relevance of the method ~~with respect to relevant legislative requirements~~ should be indicated. In particular, the proposer should indicate that the ~~principle~~ **principal** conditions for the method are fulfilled.

Scientific basis

An overview of the principles of how the method, such as DNA molecular biology based (e.g. for real-time PCR) information should be provided. References to relevant scientific publications are useful.

The prediction model adopted to interpret results and to make inferences must be described in complete detail.

Specification of the prediction model/mathematical model needed for the method

If the derivation of the results relies upon a mathematical relationship this must be outlined and recorded (e.g., a regression line or calibration curve obtained by other means). Instructions for the correct application of the model should be provided. These may include, depending on the method, a recommended number and range of levels to be analysed, minimum number of replicates to be included or the means to evaluate the goodness-of-fit**, as well as appropriate threshold values for the real-time PCR curves**.

Outline of the experimental design, including the details about the number of runs, samples, replicates etc. should be stated.

INFORMATION ABOUT THE METHOD OPTIMISATION

Primer pairs tested

For PCR methods, sufficient justification should be given of how and why the proposed primer pair**s, and if used probes,** ~~has~~**have** been selected**. This information should be provided** ~~, also~~ for the reference gene **as well** (should this be part of the method).

Specificity testing

Empirical results from testing the method with non-target **recombinant DNA**~~transgenic~~ events and non-**recombinant DNA**~~transgenic~~ plants should be provided. This testing should include closely related events and cases were the limits of the sensitivity are truly tested. In addition it might be appropriate to test other plants to reduce the potential for obtaining a false positive.

Stability testing

Empirical results from testing the method with different varieties should be provided in order to demonstrate, for instance, the stability of the copy number of the reference gene.

Sensitivity testing

Empirical results from testing the method at different concentrations in order to test the sensitivity of the method. Limits of detection must be defined using samples comprising of single crops only, e.g. "the LOD for Roundup Ready® soy is 0.1 % of total soy *if the product is comprised of 100 % soy*". For food products made up of multiple ingredients, the actual sensitivity will be reduced, as total extracted DNA will be derived from more than one ingredient so that the starting amount of the actual measurand will be decreased. This dilution effect will depend on how much of the target ingredient (e.g. soy) is in the food product and the total quantity of DNA derived from the other ingredients. Some ingredients will contribute a large amount of DNA, such as wheat or maize flour and eggs, while other ingredients will not contribute any DNA, such as sugar, water or highly processed oils.

LOD should be determined in terms of genome equivalents for each PCR system separately.

Robustness testing

Empirical results from testing the method against small but deliberate variations in method parameters.

Cross-reactivity

The cross-reactivity, interferences and matrix effects should be evaluated~~, particularly~~ for **both DNA-based and** ~~the~~ protein-based methods of analysis.

PRACTICAL APPLICATION OF THE METHOD

Applicability

Indication of the matrix (e.g., processed food, raw materials, etc.), the type of samples (e.g., seeds, flour, pizza, cookies, etc.) and the range to which the method can be applied. Relevant limitations of the method should also be addressed (e.g. inference by other analytes or inapplicability to certain situations). Limitations may also include possible restrictions due to the costs, equipment or specific and non-specific risks implied for either the operator and/or the environment.

Operational characteristics and practicability of the method

The required equipment for the application of the method should be clearly stated, with regards to the analysis *per se* and the sample preparation. An indication of costs, timing, practical difficulties, and of any other factor that could be of importance for the operators should be also indicated.

Operator skills requirements

A description of the practical skills necessary to properly apply the proposed method should be provided.

ANALYTICAL CONTROLS

The proper use of controls when applying the method should be indicated. Controls should be clearly specified and their interpretation recorded. These may include positive and negative controls, their detailed contents, the extent into which they should be used and the interpretation of the obtained values.

In particular the following should be stated:

- Positive and negative controls used
- Control samples, plasmids and alike used
- Reference materials used.

METHOD VALIDATION/PERFORMANCE

See the Codex "Check-list" (i.e. accuracy, applicability (matrix, concentration range and preference given to 'general' methods), detection limit, determination limit, precision, recovery, selectivity, sensitivity and linearity),

and in particular the following additional information should be supplied for DNA-based procedures:

- ***amplicon length***

The boundaries of the amplified product are formed by the primers at both sides. Therefore the selection of suitable primers is a crucial factor in the PCR analysis. The length of the amplified product does have a direct influence of the PCR performance. By increasing the product length, the PCR efficiency will decrease reciprocal as illustrated below (Fig. 1). In theory in every cycle the target DNA sequence is doubled (amplification factor of 2). In reality the PCR efficiency is less than 100% resulting in a decreased amount of amplified product. Moreover food processing will lead to a degradation of target DNA. Therefore the selection of shorter amplicon sizes (within reason) will increase the possibility to get a positive signal in the analysis of highly processed foodstuffs.

Amplification factor	2	1.65	1.60	1.55	1.50	1.45
Target copies after:						
10 cycles	**10^3**	**150**	**110**	**80**	**58**	**41**
20 cycles	**10^6**	**$2.2x\ 10^4$**	**$1.2x10^4$**	**$6.4x10^3$**	**$3.3x10^3$**	**$1.7x10^3$**
30 cycles	**10^9**	**$3.3x10^6$**	**$1.3x10^6$**	**$5x10^5$**	**$1.9x10^5$**	**$7x10^4$**
40 cycles	**10^{12}**	**$5x10^8$**	**$1.5x10^8$**	**$4.1x10^7$**	**$1.1x10^7$**	**$2.8x10^6$**
	100%	**82.5%**	**80%**	**77.5%**	**75%**	**72.5%**

Figure 1 PCR efficiency. A decrease of efficiency in PCR leads to lower amounts of amplified products being present after a certain number of cycles.

- ***whether the method is instrument specific***

At the moment a number of different types of real time instruments are available. These instruments may have different heating and cooling characteristics, which affects ramp rates and ~~affects~~ the time necessary for a whole PCR run. ***It is preferable for methods to be easily adapted to a number of different models and types of equipment.***

*Beside**s** the differences in the heating and cooling system**s** there are differences in the technique used to induce and subsequently to record the fluorescence. Some real time instruments use laser technique for inducing fluorescence, others are equipped only with a white lamp and filters for selecting a specific wave length. The detection of the fluorescence could also vary.*

Taking all the differences into account it is ~~impossible~~ ***often difficult*** *to change the instrument without adaptation of the PCR method. Thus,* ***the method developer should attempt to develop methods which are robust and can be performed on a range of instrument types.*** *~~because the methods are generally instrument dependent they cannot be transferred to other equipment without evaluation and/or modification.~~*

~~This is in many ways equivalent to the Codex Type I method and should be considered in the same light.~~

- ***whether single- or multi-plex PCR amplifications are undertaken***

Using more than one primer set in a single reaction is called multi-plex PCR. The aim of using such approach is to reduce costs and time for the analysis of different targets of a single sample (i.e. a ~~GMO~~ specific system is combined with a target taxon specific for relative quantitation). It must be emphasised that the unbalanced presence of one of the target sequences will lead in a preferred amplification by the polymerase during PCR. Moreover the combination of different primer sets is limited up to 7 to 10 in a single reaction.

The information provided should demonstrate the robustness of the method for inter-laboratory transferability. This means that the method should have been tested by at least one external laboratory besides the laboratory which has developed the method. This is an important pre-condition for the success of the validation of the method.

And for both protein and DNA based methods:

- ***whether there are differences between PCR-based and immunological methods concerning validation criteria***

The DNA and protein-based techniques used to detect and quantify ~~a GMO derived material in foods~~**foods derived from modern biotechnology** are based on different principles. In PCR the targeted DNA is amplified in a exponential manner, in which a small difference in the beginning of the PCR process will lead to a big difference in the amplified amount of DNA after 35-45 cycles. In contrast to that immunological detection assays are based on the direct interaction with the target molecule and do not include an amplification step.

Moreover, the quantitation by real time PCR is often based on two independent PCR systems: one for the genetic modification and one for the taxon specific sequence**, and uncertainty/methods characteristics of each must be considered**.

ANNEX II: CODEX DEFINITIONS APPLICABLE TO THE ~~GMO~~ ANALYSIS FOR FOODS DERIVED FROM MODERN BIOTECHNOLOGY

This Annex is concerned with the definitions needed in **the** ~~GMO~~ analysis **for foods derived from modern biotechnology** . ***(Note: a number of definitions have been grouped together in one heading; these may be contradictory and this needs to be resolved. The Codex definition given in the Procedural Manual should be used and amplified as necessary. Codex definitions have not been reproduced here if they need no further qualification*** ~~***for GMO analysis***~~***)***.

Accuracy

The closeness of agreement between a reported result and the accepted reference value[2].

Applicability

The analytes, matrices and concentrations for which a method of analysis may be used[3].

The analytes, matrices, and concentrations should be appropriate for the control purposes for which the method has been proposed. The description may also include warnings to known inferences by other analytes, or inapplicability to certain matrices and situations.

It is not feasible to provide reference materials for every one of the many food matrices that are available, so that the use of a representative matrix reference will usually be necessary. The use of the method in a new matrix will need to be validated at a minimum via Single Laboratory validation, usually by spike and recovery experiments, and the reference material used should be described on the report to the customer.

Dynamic Range - Range Of Quantification

The interval of concentration within which the analytical procedure has been demonstrated by collaborative trial to have a suitable level of precision and accuracy.

Limit of Detection (LOD)

Limit of detection is the lowest concentration or content of the analytes that can be detected reliably, but not necessarily quantified, as demonstrated by collaborative trial or single-laboratory validation[4]. LOD is generally expressed as the amount of analyte at which the analytical method detects the presence of the analyte at least 95% of the time ($\leq$5% false negative results).

Limit of Quantification (LOQ)

The limit of quantification of an analytical procedure is the lowest amount or concentration of analyte in a sample, which can be quantitatively determined with an acceptable level of precision and accuracy as demonstrated by satisfactory collaborative trial or single-laboratory[5] validation[6].

Practicability

The ease of operations, in terms of sample throughput and costs, to achieve the required performance criteria and thereby meet the specified purpose[7].

Generally, the method should be practical for its intended purposes.

---

[2] Definition adopted from ISO 3534-1.

[3] Slightly modified from the definition provided in Codex CX/MAS 02/4: Proposed draft guidelines for evaluating acceptable methods of analysis. Version November 2002.

[4] Slightly modified from prEN ISO 24276:2002 (E).

[5] E.g. Thompson et al. 2002. IUPAC Technical Report: Harmonised guidelines fro single-laboratory validation of methods of analysis. Pure Appl. Chem. 74(5): 835-855.

[6] Slightly modified from prEN ISO 24276:2002 (E).

[7] Adopted from prEN ISO 24276:2002 (E).

Repeatability standard deviation (RSD_r)

The standard deviation of test results obtained under repeatability conditions. Repeatability conditions are conditions where test results are obtained with the same method on identical test items in the same laboratory by the same operator using the same equipment within short intervals of time.[8]

Reproducibility standard deviation (RSD_R)

The standard deviation of test results obtained under reproducibility conditions. Reproducibility conditions are conditions where test results are obtained with the same method on identical test items in different laboratories with different operators using different equipment.[9]

Recovery

Proportion of the amount of analyte, present in or added to the analytical portion of the test material, which is extracted and presented for measurement.

Ruggedness (Robustness)

Robustness refers to variations in the method as performed in different laboratories by different technicians. The language used here is derived from "Ruggedness" which is the equivalent in the harmonized guidelines. Ruggedness should be demonstrated by the validation of the method in 8-12 laboratories as defined in the harmonized guidelines. It is preferable from a CODEX point of view, that these laboratories be distributed across several continent/trading blocks.

The robustness of an analytical method is a measure of its capacity to remain unaffected by small, but deliberate variations in method parameters and provides an indication of its reliability during normal usage[10].

Sensitivity

The sensitivity of a method is a measure of the magnitude of the response caused by a certain amount of analyte.

The method should be sensitive enough in order to be able to detect/quantify with respect to the thresholds as provided in the relevant legislation.

Since sensitivity is method- and purpose-dependent it should be specified in the protocol. A reasonable goal for sensitivity is that required to meet levels specified in contracts, with a reasonable certainty that the level does not exceed the required limit.

Sensitivity as a term is used in two different ways - LOD and the slope of a curve. The use of "detection limit", or "limit of detection" is the preferred term to use as a measure of the ability of a method to detect a small amount of analyte. See also previous comments regarding sensitivity in this document.

Specificity

Property of a method to respond exclusively to the characteristic or analyte of interest.

Trueness

The closeness of agreement between the average value obtained from a large series of test results and an accepted reference value[11].

[8] Definitions adopted from ISO 3534-1.

[9] Definitions adopted from ISO 3534-1

[10] Definition adopted from ICH Topic Q 2 A "Validation of analytical methods: definitions and terminology." The European Agency for the evaluation of medicinal products. CPMP/ICH/381/95. Version November 1994. http://www.emea.eu.int/pdfs/human/ich/038195en.pdf

[11] Adopted from ISO 3534.

The measure of trueness is usually expressed in terms of bias. Trueness has also been referred to as "accuracy of the mean".

ANNEX III: VALIDATION OF A QUANTITATIVE PCR METHOD

INTRODUCTION

DNA-based analysis is commonly performed using Polymerase Chain Reaction (PCR). This technique amplifies a specific (short) segment of DNA to the extent that its quantity can be measured instrumentally (e.g., using fluorometric means). As DNA is a molecule that is easily degraded during food processing operations (e.g., due to heat, enzymes and mechanical shearing), we urge that this be considered in the performance criteria assessment of this technique. This is relevant as in most foods raw ingredients are not present, but are in a processed form, which has an effect on proteins and/or DNA present in food. Furthermore, these protein(s) and/or DNA may be degraded, or its total amount may be decreased due to processing. As a result, any current detection method (DNA- or protein-based) is affected.

It is often the case that the results of a determination are expressed in terms of percent of a sample that contains a particular biotechnology-derived sequence. In a quantitative test, this measurement actually involves two PCR-based determinations – that of the primary analyte (e.g. an inserted gene sequence) and that of the endogenous, or comparator sequence (e.g. an endogenous maize gene). Each of these determinations has its own uncertainties, and the two are likely to have different measurement characteristics. In most applications, the primary analyte will be present at low concentrations, and the comparator will be present at concentrations 10 to 1000 times higher. It is thus important that both measurements are properly validated. In cases where the measurement is expressed directly as a percentage (as in the use of ΔCT), these factors must be considered when validating the method.

The consequence is that the analysis of DNA, especially in processed foods, aims at detecting a very small amount of analyte. Although the result of a PCR analysis is often expressed in % as the relative amount of DNA specific for foods derived from modern biotechnology relative to the total amount of DNA for a specific species, the actual amount of DNA specific for foods derived from modern biotechnology is often in the nanogram/gram range or lower. Analysis of those low amounts of analyte is accompanied by a considerable measurement uncertainty; this needs to be appreciated by the users of analytical results.

VALIDATION

A quantitative PCR assay should be validated for the intended use or application. A harmonized ISO/IUPAC/AOAC protocol was developed for chemical analytical methods. This defines the procedures necessary to validate a method (Horwitz W; Protocol for the design, conduct and interpretation of method-performance studies. Pure and Applied Chemistry, 67, 331 (1995)). It is important to emphasize that all the principles and rules of the harmonized protocol are applicable to quantitative PCR methods.

A number of the parameters involved in validation of the performance of a quantitative PCR assay will be discussed in detail. These are scope, LOD and LOQ, accuracy, precision, sensitivity and ruggedness (robustness). Other important factors are acceptance criteria and interpretation of results, and the issue of the units in which results are expressed.

It is important to note that a quantitative PCR assay typically consists of two assays, one determines the amount of DNA specific for the **recombinant DNA**~~transgenic~~ product, while the other is specific for the amount of plant specific DNA. Each of these assays has to be considered separately, as these assays can be considered as independent analytical procedures. Thus, all parameters listed below, including specificity and sensitivity, have to be assessed individually for each of the assays involved. These are given alphabetically, not necessarily in order of importance.

Accuracy

As for any method, the accuracy of a method should be compared to known values derived from reference materials, ideally the best characterised. Precision will be determined in the usual way from single laboratory (repeatability) and multi-laboratory (reproducibility) studies.

Recommendation: The accuracy should be within [± 35%] of the accepted reference value over the whole dynamic range.

Applicability

The analytes, matrices and concentrations for which a method of analysis may be used must be stated.

Dynamic Range - Range Of Quantification

The scope of the methods defines the concentration range over which the analyte will be determined. Typically the range for a ~~GM~~ **recombinant DNA** product will range from a tenth of a percent up to a few percent and for the endogenous control the range will be close to 100%, unless the testing of complex mixtures is envisioned. This desired concentration range defines the standard curves and a sufficient number of standards must be used to adequately define the relationship between concentration and response. The relationship between response and concentration should be demonstrated to be continuous, reproducible and should be linear after suitable transformation.

The range of a quantitative method is typically designed to be in the range 0.1% - 100% (DNA %w/w). However, it is common to validate a method for a range of concentrations that is relevant to the scope of the application. If a method is validated for a given range of values, the range may not be extended without validation. For certain applications (e.g., seed or grain analysis) the use of genomic DNA for the preparation of the standard curve (see discussion on the use of plasmid DNA below) may be considered. While it is easy to establish a nominal 100% standard (limited only by the purity of the materials used) it is difficult to reliably produce standard solutions below 0.1%. This is due to the uncertainties involved in measuring small volumes and the error propagation if serial dilution steps are applied. Additionally, the number of target sites (DNA sequence to be amplified) becomes so small that stochastic errors will begin to dominate and no reliable analysis is possible[1, 2]. If genomic DNA is chosen to be used as calibrator, it is important that this calibrator needs to be traced back (in its metrological meaning) to a reference of highest metrological order, e.g. a certified reference material. The range will be established by confirming that the PCR procedure provides an acceptable degree of linearity and accuracy when applied to samples containing amounts of analyte within or at the extremes of the specified range of the procedure.

Recommendation: The dynamic range of the method should cover at least [20% and 5] times the target concentration, where practicable. ~~*Target concentration should be understood here as the threshold relevant for a certain regulation.*~~

*Example: 0.~~2~~**1**% and ~~2~~**5**.0% for a 1% ~~GMO~~ **target** concentration or ~~50~~ **100** and ~~1000~~ **2500** genome copies if the target is 500 copies.*

There is a general scientific discussion still going on about the interpretation of the percentage values (e.g. dynamic range from **2**~~1~~0% to 5 times the target value). Although the experts agreed that – at least for PCR – copy number is desired over weight/weight percentage, it was recognised that so far there is no reliable weight/copy number relationship because of inter-variety fluctuation of the 1C value and because of uncertainty in the correlation of weight of ingredient to weight of DNA. ~~For the time being, both the w/w and copy number/copy number calculations are acceptable.~~

The unique characteristics of quantitative PCR impose particular restrictions on the low end of the dynamic range of a quantitative PCR. This is due to the difficulty in determining LOD and LOQ values due to the non-normal distribution of variances in the values in this range. Thus it may not be appropriate to require a range extending to **2**~~1~~0% of the measured value. The suggestion of a dynamic range that ranges from 10 to 200% can be problematic. For example, capability to analyse a foodstuff composed of more than 50% (w/w) of a biotechnology-derived material (as might be the case for a nutraceutical) would require a dynamic range exceeding 100% (w/w). This is clearly not possible.

Limit of Detection (LOD)

If the validation of the quantitative PCR assay shows that the assay can measure **recombinant DNA**~~transgenic~~ plant DNA at (for example) 0.1% with acceptable trueness and precision, then it is often not necessary to determine the LOD and LOQ, as the method is only being applied above the range where these are relevant. However, if the method is being used at concentrations close to the limit of detection and limit of quantification (typically 0.01-0.05%), then the assessment of the LOD and LOQ will become part of the validation procedure.

It is worth noting that a determination of an LOD or LOQ is not necessarily needed to establish the validity of a method for a given application. For example, it does not add much value if an LOD is determined to be 1 ng/kg, while the scope of the method validation extends only for concentrations ranging in g/kg. In this and similar cases the reliability of the method will be proven by the other parameters and no efforts are included in the method validation to assess the LOD. However, the LOQ shall always be established and included in the validation study **for a quantitative method**.

If the LOD is required, it is common practice to assume that it is the signal strength of a blank increased by three times the standard deviation of the blank. However, this method gives at best an estimate, relies on normal Gaussian distribution of the blank measurements around zero, and may give a lower value than the actual LOD. Its use is not valid in methods such as ~~Q~~**q**uantitative PCR, in which the distribution of measurement values for blanks is typically truncated at zero and is thus not normally distributed. Thus the LOD need to be experimentally determined unless the targeted concentrations are well above the LOD and the LOD therefore becomes irrelevant. For **qualitative**~~quantitative~~ methods the LOD is the amount of analyte at which the analytical method detects the presence of the analyte at least 95% of the time (≤5% false negative results). This, and the false positive rate, are the only parameters required for a qualitative method other than specificity.

For a quantitative method, it is important to know whether the LOQ for a particular matrix is close to the values to be measured. Using the traditional approach, the LOQ can be expressed as the signal strength of a blank equal to the LOD increased by 6-10 times the standard deviation of the blank, unless it is known from other sources that the measured values range so high above the LOQ that its knowledge becomes irrelevant. However, this method to determine the LOQ leads only to an estimate of the true LOQ that may be an artificially high or low approximation.

In practice, two procedures have been employed to determine the LOQ. The first approach is to assay a number of negative samples that have been supplemented (spiked) with known amounts of analyte. The LOQ is then the level at which the variability of the result and percent recovery of the analyte meet certain preset criteria. For small molecules, these criteria have typically been a CV of <20% and 70-120% recovery. DNA recovery, however, may be difficult from some matrices, e.g. starches or ketchup, and lower recovery efficiencies may have to be accepted. When recovery efficiencies are low, this must be stated in the validation data and in the analytical report. A more complete approach is to test the method using a number of samples that contain known amounts of the ~~GM~~ **recombinant DNA** material. This is more complicated as it requires access to significant quantities of reference materials that contain a known range of concentrations of the ~~GM~~ **recombinant DNA** event of interest. Procedures for assessing LOD and LOQ during the validation of quantitative PCR methods are discussed in Annexes III and IV.~~.~~

Validation of methods consists of two phases. The first is an in-house validation of all of the parameters above except reproducibility. The second is a collaborative trial, the main outcome of which is a measure of the repeatability and reproducibility together with detailed information on the transferability of methods between laboratories. It is strongly recommended that a small-scale collaborative trial be performed to test the general ruggedness of a particular method before the expense of organizing a large-scale trial is incurred. In case any improvement of the method or the method description are needed, only limited expenses are incurred through the pre-trial, while a failure of a full interlaboratory method validation due to a ambiguous method description is a very costly failure. Additionally, it may be pointed out that the implementation of an already validated method in a laboratory needs to include necessary experiments to confirm that the implemented method performs as well under local conditions as it did in the interlaboratory method validation. It is important to note that a method should be validated using the conditions under which it will be performed.

Recommendation: Limit of detection is to be < 10% of the ***target*** *value~~ of specification~~. ~~The value of specification should be understood here as the threshold relevant for a certain application~~.*

Note: limits of detection must be defined using samples comprising of single crops only, e.g. "the LOD for Roundup Ready® soy is 0.1 % of total soy if the product is comprised of 100 % soy". For food products made up of multiple ingredients, the actual sensitivity will be reduced, as total extracted DNA will be derived from more than one ingredient so that the starting amount of the actual measurand will be decreased. This dilution effect will depend on how much of the target ingredient (e.g. soy) is in the food product and the total quantity of DNA derived from the other ingredients. Some ingredients will contribute much DNA, such as wheat or maize flour and eggs, while other ingredients will not contribute any DNA, such as sugar, water or highly processed oils.

Limit of Quantification (LOQ)

See introduction above for limit of detection.

Recommendation: The limit of quantification is to be < 20% of the value of specification with an RSD_r ≤ [25%] or as close as is practicable. ~~*The value of specification should be understood here as the threshold relevant for a certain regulation*~~

Example: For a 1 % nominal value LOQmin = 0.1 % or for 500 copies LOQmin = 50 copies.

For a quantitative method, it is important to know whether the LOQ for a particular matrix is close to the values to be measured. Traditional methods of approximating the LOQ (zero value plus 6-10 standard deviations) rely on normal Gaussian distribution of the blank measurements around zero. This approach is not valid in methods such as Quantitative PCR, in which the distribution of measurement values for blanks is typically truncated at zero and is thus not normally distributed. Thus the LOQ needs to be experimentally determined.

Practicability

The practicability of the method must be demonstrated.

Repeatability standard deviation (RSD_r)

Recommendation: The relative repeatability standard deviation should be below 25% over the whole dynamic range of the method.

Reproducibility standard deviation (RSD_R)

Recommendation: The relative reproducibility standard deviation should be below 35% at the target concentration and over the majority of the dynamic range. RSD_R < 50% at the limit of quantification/lower end of the dynamic range.

Ruggedness (Robustness)

The evaluation of ruggedness (robustness) demonstrates the reliability of a method with respect to inadvertent variation in assay parameters. Variations that may be included are reaction volumes (e.g., 25 vs. 30µl), annealing temperature (e.g., plus and minus 1°C) and/or other relevant variations. The experiments need to be performed at least in triplicates and the recovery needs to be calculated. The response of an assay with respect to these small changes should not deviate more than ±30% from the response obtained under the original conditions.

Note: Recovery may be difficult to determine.

The adequacy of the robustness testing needs to be analysed on method-by-method basis. For instance, for a real-time PCR method, the following factors should ideally be taken into account: different thermal cycler models, DNA polymerase, uracil-n-glycosylase, magnesium chloride concentration, primer forward and reverse concentration, probe concentration, temperature profile, time profile, dNTP including dUTP concentrations.

Sensitivity

For a quantitative PCR method, a linear relationship of the CT as a function of the logarithm of the concentration of the target of the individual target should be obtained across the range of the method. The correlation coefficient, y-intercept, slope of the regression line and % of residual should be reported. The % of residual for each of the calibrators should preferably be ≤=30%.

In order to obtain a standard curve for event specific quantitative assays, standard DNA mixtures can be prepared by combining purified genomic DNA from **recombinant DNA**~~transgenic~~ and non- **recombinant DNA**~~transgenic~~ plants material such as seed or leaves. The content of **recombinant DNA**~~transgenic~~ plant DNA in the mixtures might be 100, 50, 10, 5, 1, 0.5, 0.1, and 0% or as appropriate for a smaller concentration range. Three replicates must be analysed for each point on the standard curve. **Alternatively the 5% CRM may be diluted without target DNA.**

For quantitative assays on plant endogenous genes, standard DNA mixtures can be prepared by combining purified genomic DNA from the target plant species and that of a non-target plant species. For example, for validation of a maize ADH1 quantitative assay, the target plant species is maize and the non-target plant species could be soybean or another species. The content of DNA of the target plant species in the mixtures is typically 100, 50, 25, 10, 5, 1 and 0% or as appropriate. Three replicates must be analysed for each point on the standard curve. ~~Alternatively the 5% CRM used and further diluted without target DNA.~~

In cases where the ΔCT-method is employed, it will be the responsibility of the analyst to ensure that the overall amount of DNA is well within the range for which the assay was validated.

~~*Recommendation: Typical sensitivities are in the range of 0.1% biotechnology-derived material by weight if the material is not highly processed.*~~

Specificity

The specificity should be demonstrated by showing experimental results from testing the method with non-target **recombinant DNA**~~transgenic~~ events and non- **recombinant DNA**~~transgenic~~ plants. This testing should include closely related events and cases where the limits of the detection are truly tested. As the method should be event-specific it should only be functional with the ~~GMO or GM based~~ **recombinant DNA** product considered and ought not to be functional if applied to other events already authorised. In addition, if a reference gene system is a part of the method this should not recognize any gene corresponding to even phylogenetically related species, and should give similar CT-values when amplifying equal amounts of DNA from different cultivars of the same species.

The adequacy of the testing needs to be analysed on a method-by-method basis. It will be necessary to obtain information about the specificity testing in case of stacked genes at some stage.

*Recommendation: Specificity is the starting point for a method and needs to be considered during primer **and probe** design. Primers **and probes** should be checked against the known sequence of the event insert and pertinent databases for possible matches. Specificity must also be demonstrated experimentally. The following suggests a reasonable approach and the experiments should be performed during pre-validation of an assay.*

For event-specific assays:

- *Analyse at least a total of ten non-target **recombinant DNA**~~transgenic~~ events and any non- **recombinant DNA**~~transgenic~~ plants that may commonly be found as contaminants in the commodity.*
- *Test on sample from each source (total of at least 10 DNA samples).*
- *Analyse two replicates for each DNA sample.*

Test results shall clearly indicate that no significant instrument reading is observed.

For assays on plant endogenous genes:

- *Analyse at least a total of ten different plant samples comprising different varieties of the same plant species as well as other plants species important for food production (such as wheat, rice, corn, potato, and soybean) and that may commonly be found as contaminants in the commodity.*
- *Test one sample from each source (total of at least 10 DNA samples).*
- *Analyse two replicates for each DNA sample.*

Test results shall clearly indicate that no significant instrument reading is observed.

Trueness

Recommendation: The trueness should be within ± [30%] of the accepted reference value over the whole dynamic range. This refers to the PCR-step provided that a modular validation approach has been applied.

ANALYTICAL CONTROL ACCEPTANCE CRITERIA AND INTERPRETATION OF RESULTS

A validated method also includes criteria on which the validity of an observed measurement result can be accepted as valid. It is important to follow these criteria and to observe the rules for data interpretation. In the case that it may be desired to deviate from said criteria and rules a new method validation study would be needed in order to demonstrate the validity of the new rules and procedures.

At a minimum, the following acceptance criteria are common to all quantitative PCR methods and applicable to each PCR run:

- The result of the positive DNA target control, with, for example 1% **recombinant DNA**~~transgenic~~ DNA, the mean of the replicates deviates less than 3 standard deviations from the assigned value. A target DNA control is defined as reference DNA or DNA extracted from a certified reference material or known positive sample representative of the sequence or organism under study. The control is intended to demonstrate what the result of analyses of test samples containing the target sequence will be.

- The amplification reagent control is $\leq$ LOD. The amplification control is defined as control containing all the reagents, except extracted test sample template DNA. Instead of the template DNA, a corresponding volume of nucleic acid free water is added to the reaction.

- The % of residual for each of the standards should be <30%

To accept the result of an unknown sample, the relative standard deviation of the sample replicates should be <[35]%.

REFERENCES FOR ANNEX III

1. Huebner P, Waiblinger H U, Pietsch K, Bordmann P (2001) Validation of PCR methods for quantitation of genetically modified plants in food. Journal of AOAC International 84(6) 1855-1864;

2. Kay S, Van den Eede G, The limits of GMO detection, Nature Biotech. 19(5) 504 (2001)).

3. Residue Chemistry Test Guidelines OPPTS 860.1340 "Residue Analytical Method" United States Environmental Protection Agency, August 1996, (Mihaliak & Berberich, 1995).

ANNEX IV: VALIDATION OF A QUALITATIVE PCR METHOD

Introduction

A qualitative PCR must be validated in the same way as it is intended to be used – that is the sensitivity of the method must be shown to be such that it can reliably detect one positive particle (seed) in a pool, and does not give rise to a significant number of false positives. A concept of using false-positive and false-negative rates to describe the accuracy and precision of a qualitative assay has been developed for microbial assays[1]. This concept can be applied to qualitative PCR assays. A critical issue in the validation of this type of method is the availability of test materials that are known to be positive and negative. The provision of negative reference materials is particularly important and critical in the case of a qualitative method. Any impurities must be present only at levels so low that they become negligible.

By their very nature, qualitative test results refer to the identification above/below a limit. The measures of precision and accuracy are the frequencies of false negative and/or false positive results at the detection limit. False negative results indicate the absence of a given analyte when in fact the analyte is present in the sample, while false positive results indicate the presence of an analyte that is not present in the sample. Due to the inherent nature of the analytical technique, an increase in false negative results will be observed when the amount of analyte approaches the LOD of the method. Like the limit of detection for quantitative methods, the limit of detection for a qualitative method can be defined as the concentration at which a positive sample yields a positive result at least 95% of the time. This results in a rate of false negative results of 5% or less. During validation of a qualitative PCR assay, it is also important to determine the number of false positive results (-a positive result obtained using a sample that is known to be negative). This is also expressed as a rate.

False Positive Rate

This is the probability that a known negative test sample has been classified as positive by the method. The false positive rate is the number of misclassified known negatives divided by the total number of negative test samples (misclassified positives plus the number of correctly classified known negatives) obtained with the method:

For convenience this rate can be expressed as percentage:

$$\% \text{ false positive results} = \frac{\text{number of misclassified known negative samples}}{\text{total number of negative test results [incl. misclassified]}}$$

Note: different sectors use different definitions here.

False Negative Rate

This is the probability that a known positive test sample has been classified as negative by the method. The false negative rate is the number of misclassified known positives divided by the total number of positive test samples (misclassified positives plus the number of correctly classified known positives) obtained with the method.

For convenience this rate can be expressed as percentage:

$$\% \text{ false negative results} = \frac{\text{number of misclassified known negative samples}}{\text{total number of positive test results [incl. misclassified]}}$$

Note: different sectors use different definitions here.

In order to demonstrate the false negative rate for qualitative assay, a series of samples (e.g. grain/seed pools) with a constant, known concentration of positive material in a pool of negative material (e.g., 1 positive kernel in 199 conventional corn kernels) have to be analysed and the results evaluated. It is important to note that the concept of confidence intervals and statistical uncertainty needs to be applied to the risk of false positive and/or false negative results as well. The desired level of confidence determines the size and number of pools that need to be tested. For example, 100 positive test results obtained from 100 independent

measurements on truly positive samples lead to the conclusion that the level of false negative results is below 4.5% at a confidence level of 99% for the tested concentration of positive kernels (expressed as the number of positive kernels in a pool of negative kernels).

Ruggedness

As with any validated method, reasonable efforts must be made to demonstrate the ruggedness of the assay. This involves careful optimisation and investigation of the impact of small modifications that could occur for technical reasons.

Acceptance Criteria and Interpretation of Results

A validated method includes criteria on which the validity of an observed measurement result can be accepted as valid. It is important to follow these criteria and to observe the rules for data interpretation. It is therefore important to make sure that the result of the positive DNA target control, is positive. Similarly the amplification reagent control must be negative. In addition to these controls, it is desirable to carry out a parallel reaction on the same DNA sample using a primer set which detects an endogenous single copy sequence. This reaction is carried out on every DNA sample, and can either be in the same reaction (multiplexed) or as a separate reaction. In the case of multiplexed reactions, it is important that the endogenous reaction does not out compete the event specific reaction for reagents, as the endogenous sequence is likely to be present at up to 1000 fold the amount of the target sequence. The control reaction with the endogenous sequence gives an indication of the quality of the DNA as a template for the PCR reaction. Table 1 sets out the accept/reject criteria for the PCR reactions on a per lane basis, using the results of the PCR reaction with the endogenous sequence.

Table 1: Criteria for scoring Qualitative PCR analyses

PCR result (~~GM~~ analyte)	PCR result (endogenous)	Scoring of test
+	+	+
-	+	-
+	-	[+][-]
-	-	Reject

A further complication is however introduced by the fact that qualitative PCR reactions are typically carried out in duplicate. Thus it can occur that the duplicates do not agree. It is common practice to repeat PCR reactions once on DNA samples that are rejected. A repeated indeterminate result is indicative that the analyte cannot be reliably detected. (Table 2), and that the assay is operating below the limit of detection as, by definition, a 95% or better detection rate would be achieved at the limit of detection. The sample is therefore scored negative. Similar criteria apply if more replicates are carried out on each DNA sample.

Table 2: Criteria for scoring duplicate qualitative PCR analyses

Lane 1	Lane 2	Scoring of test
+	+	Positive
-	+	Repeat/Indeterminate
+	-	Repeat/Indeterminate
-	-	Negative

REFERENCES FOR ANNEX IV

1. AOAC® Official Methods℠ Program Manual, Appendix X p14f, May 2002, AOAC International; http: www.aoac.org/vmeth/omamanual/htm.

ANNEX V: VALIDATION OF A PROTEIN-BASED METHOD[1]

QUANTITATIVE TESTING

Quantitative immunoassays are used to determine levels of the protein analyte in specific parts of the plant (e.g. seed, leaf, root, stalk etc). Typical applications are given in Table 1. In order to perform a microplate ELISA for quantitative determination of a protein analyte in plant tissue, it is first necessary to obtain a representative sample of the plant material. The sample amount will influence the detection limit or sensitivity of the assay. The analyte is then extracted from the plant material by adding a solvent and blending, agitating, or applying sheering or sonic forces. Typical solvents used are water or buffered salt solutions. Sometimes detergents or surfactants are added. Some proteins require more rigorous procedures like homogenisation or boiling in solvents, detergents, salts etc.

After the capture antibody has been immobilized on the microplate well surface, a precise volume of the standard or sample extract solution is added to each well. The analyte in the test solution binds to the capture antibody. The enzyme-labelled second antibody is then added and also binds to the analyte, forming a sandwich. At this point, the well is washed to remove unbound analyte and antibodies, leaving only the antibody-analyte-antibody complex bound to the well surface. A colorimetric substrate is added which reacts with the enzyme label and produces a coloured product. The reaction is stopped after a set period of time and the colour absorbance at a given wavelength is measured on a photometer. The standard curve is generated by plotting the optical density (OD) on the *y*-axis (linear scale) against the concentration on the *x*-axis (log scale) which produces a sigmoidal dose response curve Figure 4.

To obtain an accurate and precise quantitative value, the OD for the sample solutions must fall on the linear portion of the standard curve. If the OD is too high, the sample solution must be diluted until the OD falls within the quantitative range of the assay. The concentration of the protein analyte in the original sample of plant material is calculated by correcting for any dilution factor that was introduced in preparing the sample for application to the microplate. The initial weight of the sample and the volume of extraction solvent, as well as any subsequent dilutions are used to calculate the dilution factor.

Various assay controls can be employed to demonstrate the performance of the assay. A blank sample such as an empty well or buffered solution can be run in the assay to determine any background response which can be subtracted from sample and standard responses if desired. A negative control sample (i.e. matrix extract solution known to contain no analyte) can be used to demonstrate whether a non-specific response or matrix effect is occurring in the assay. A positive control or matrix extract spiked with a known amount of the analyte can be run to demonstrate accuracy. Standards and samples can be run in replicate to demonstrate precision. Blanks, negative controls, positive controls, fortified sample extracts, standardized reference material extracts, and replicates are typically run on each microplate to control for plate-plate variation.

STANDARDIZED REFERENCE MATERIALS

The standardized reference material consists of the same matrix as the actual agricultural commodity to be tested. For example, if the matrix to be tested is soybean seed, the standardized reference material would be soybean seed containing a known proportion of **recombinant DNA**~~transgenic~~ seed. Alternatively, a pure sample or extract of the protein of interest may be used, providing the use of such protein reference materials has been validated against the matrix in question. Access to standardized reference materials is important during the development, validation, and use of immunoassays for analysis of introduced proteins in **recombinant DNA**~~transgenic~~ agricultural commodities. The best available reference material should be used in order to comply with regulations and testing needs.

In the case of commodities such as grain or seed, where the commodity consists of discrete units, it is fairly straightforward to make a reference sample with a known proportion of **recombinant DNA**~~transgenic~~ material. In other cases, generating reference samples for certain matrices and analytes can be difficult. Stability and uniformity are important considerations. For example, if the matrix to be tested consists of a mixture of materials, it would be difficult to combine **recombinant DNA**~~transgenic~~ and non- **recombinant DNA**~~transgenic~~ material in such a way as to achieve a homogeneous reference sample with a known proportion of **recombinant DNA**~~transgenic~~ material. The stability of these materials would need to be

evaluated under storage and test conditions. In any case, it is useful to have non- **recombinant DNA**~~transgenic~~ and **recombinant DNA**~~transgenic~~ material available to use as negative and positive controls.

During assay development, the reference material is used to help select assay parameters which would minimize any interfering effects of the matrix (e.g. non-specific binding of sample components to the antibodies). During validation and use of the assay, the reference materials can be extracted and analysed alongside the test samples so that the results can be directly compared.

VALIDATION OF A QUANTITATIVE PROTEIN-BASED METHOD

The principles of method validation described in appendices III and IV for PCR methods also apply to protein methods. For commercially available immunoassay kits, assay performance is generally validated by the manufacturer and is documented in the product user's guide.

Quantitative protein-based methods are better characterized as a class than PCR-based methods. Validation should be conducted according to the harmonized ISO/IUPAC/AOAC protocol **that** was developed for chemical analytical methods. This defines the procedures necessary to validate a method[2].

Accuracy: Accuracy is demonstrated by measuring the recovery of analyte from fortified samples and is reported as the mean recovery at several levels across the quantitative range. Ideally, quantitative methods will have demonstrated recoveries between 70 and 120% and a coefficient of variation (CV) of less than 20% for measured recoveries at each fortification level (Mihaliak & Berberich, 1995).

Extraction efficiency: Extraction efficiency is a measure of how efficient a given extraction method is at separating the protein analyte from the matrix. It is expressed as percent analyte recovered from the sample. Since the introduced protein expressed is endogenous to the plant, it can be difficult to demonstrate efficiency of the extraction procedure. There may not be an alternate detection method against which to compare the immunoassay results. One approach to addressing extraction efficiency is to demonstrate the recovery of each type of introduced protein analyte from each type of food fraction by exhaustive extraction, i.e. repeatedly extracting the sample until no more of the protein is detected (Stave, 1999).

Precision: Intra-assay precision describes how much variation occurs within an assay. **I**~~i~~t can be evaluated by determining the variation (%CV) between replicates assayed at various concentrations on the standard curve and on the pooled variation (%CV) derived from absorbance values in standards from independent assays performed on different days. Interassay precision describes how much variation occurs between separate assays and can be measured by analysis of quality control samples on every microplate. The quality control samples required would consist of two pools of extracts, one extract from **recombinant DNA**~~transgenic~~ plant tissue and one from conventional plant tissue. These extracts would be stored frozen and a portion would be thawed and assayed on every microplate. Interassay precision could be evaluated over time and expressed as % CV (Rogan *et al*, 1999). The precision of protein-based quantitative methods is in general higher than PCR-based methods.

Recommendation: The accuracy should be within [± 25%] of the accepted reference value over the whole dynamic range.

Sensitivity: The sensitivity of the assay could be defined as the amount of analyte that can be measured by an absorbance reading of two standard deviations above background absorbance (Rogan *et al*, 1992). The detection limit could be expressed as the lowest dilution of **recombinant DNA**~~transgenic~~ crop that could be detected when **recombinant DNA**~~transgenic~~ and non- **recombinant DNA**~~transgenic~~ crop are combined (Rogan *et al*, 1999).

Dynamic Range - Range Of Quantification

The scope of the methods defines the concentration range over which the analyte will be determined. In most cases the analytical range for a **recombinant DNA** ~~GM~~ product will range from a tenth of a percent up to a few percent. This desired concentration range defines the standard curves and a sufficient number of standards must be used to adequately define the relationship between concentration and response. The

relationship between response and concentration should be demonstrated to be continuous, reproducible and should be linear after suitable transformation.

Interpretation of the percentage values (e.g. dynamic range from 10% to 5 times the target value) can be difficult when using quantitative methods. Quantitative protein methods generally give an estimate of the concentration of the ~~GM~~ protein in the matrix, due to variations in the expression of the amount of protein in different tissues of plants, and within the same tissue at different locations. The use of qualitative protein-based methods is thus much more prevalent. In addition, care must be taken to employ a method which can detect the protein in the matrix. For example, it is believed that proteins undergo modification or degradation due to processing to a greater degree than DNA, and thus loss of signal due to processing effects must be considered.

It is worth noting that a determination of an LOD or LOQ is not necessarily needed to establish the validity of a method for a given application. For example, it does not add much value if an LOD is determined to be 1ng/kg, while the scope of the method validation extends only for concentrations ranging in g/kg. In this and similar cases the reliability of the method will be proven by the other parameters and no efforts are included in the method validation to assess the LOD. However, the LOQ shall always be established and included in the validation study.

Limit of Detection (LOD)

LOD is defined in annex II. **Typically, P**~~p~~roteins are present in ~~GM~~ foods **derived from modern biotechnology** at higher concentrations than the target DNA for PCR methods. Thus stochastic effects have less influence on the determination of the LOD than when using PCR.

It is common practice when estimating the LOD to assume that it is the signal strength of a blank increased by three times the standard deviation of the blank. This method gives at best an estimate, and relies on normal Gaussian distribution of the blank measurements around zero. This can generally be assumed for methods such as ELISA, but the LOD is best determined experimentally. Alternatively the LOD is commonly defined as a concentration equal to the lowest standard used in the assay, should a positive value be consistently obtained with that standard.

Limit of Quantification (LOQ)

For a quantitative method, it is important to know whether the LOQ for a particular matrix is close to the values to be measured. Using the traditional approach, the LOQ can be expressed as the signal strength of a blank equal to the LOD increased by 6-10 times the standard deviation of the blank, unless it is known from other sources that the measured values range so high above the LOQ that its knowledge becomes irrelevant. However, this method to determine the LOQ leads only to an estimate of the true LOQ that may be an artificially high or low approximation.

In practice, two procedures have been employed to determine the LOQ. The first approach is to assay a number of negative samples that have been supplemented (spiked) with known amounts of analyte. The LOQ is then the level at which the variability of the result and percent recovery of the analyte meet certain preset criteria. For small molecules, these criteria have typically been a CV of <20% and 70-120% recovery[3]. Protein recovery, however, may be difficult from some matrices, e.g. starches or oils, and lower recovery efficiencies may have to be accepted. When recovery efficiencies are low, this must be stated in the validation data and in the analytical report. A more complete approach is to test the method using a number of samples that contain known amounts of the **recombinant DNA** ~~GM~~ material. This is more complicated as it requires access to significant quantities of reference materials that contain a known range of concentrations of the **recombinant DNA** ~~GM~~ event of interest. Procedures for assessing LOD and LOQ during the validation of quantitative PCR methods are also discussed in annexes III and IV.

Specificity

The specificity is the degree to which analogs or other molecules bind to the antibodies and should be characterized and described in the method. Specificity should be demonstrated by showing experimental results from testing the method with non-target **recombinant DNA**~~transgenic~~ events and non- **recombinant**

DNA~~transgenic~~ plants. This testing should include closely related events and cases where the limits of the detection are truly tested. As the method should be protein-specific it should only be functional with the ~~GMO or GM~~ **recombinant DNA** ~~based~~ products considered and ought not to be functional if applied to events which do not express the protein in question. Interferences: the potential for interferences from reagents and labware can be evaluated by assaying extracts from non- **recombinant DNA**~~transgenic~~ plant material.

Matrix effects: **I**~~i~~f the response of the method is affected by a substance in the final extract other than the specific protein analyte, the non-specific response is referred to as a matrix effect. One way to manage matrix effects is to demonstrate that the analytical method gives identical results with or without sample matrix present in the extract. In this approach, freedom from matrix effects would have to be demonstrated in all matrices for which the assay is to be used. Another approach (although less desirable) to managing matrix effects would be to prepare the standard solutions in extracts from non- **recombinant DNA**~~transgenic~~ matrix, i.e. matrix-matched standards. This would ensure that any matrix effects would be consistent between the standards and the samples.

Ruggedness (Robustness)

The evaluation of ruggedness (robustness) demonstrates the reliability of a method with respect to inadvertent variation in assay parameters. Variations that may be included are reaction volumes incubation temperature (e.g., plus and minus 5-10^{o}C) and/or other relevant variations. The experiments need to be performed at least in triplicates and the recovery needs to be calculated. The response of an assay with respect to these small changes should not deviate more than ±30% from the response obtained under the original conditions. Experiments which may be performed to establish ruggedness include repeated analysis of a sample or samples on several days and measurement of accuracy and precision in fortified samples using control material from several sources.

QUALITATIVE (THRESHOLD) TESTING

Lateral flow devices are useful tools for on-site or field ~~threshold~~ testing. This type of testing requires a quick, accurate and cost-effective approach. In order to ensure reliable results, the manufacturer of the lateral flow device must conduct a method validation and provide a description of the performance characteristics of the product in the package insert. If this has been completed there is generally no need for validation studies to be performed by users. Each lateral flow device is an individual stand-alone unit, capable of performing to the standards described in the product package insert.

In order to establish an on-site procedure for ~~threshold~~ testing, the ~~threshold~~**target** level must first be established. To establish that the lateral flow device is able to differentiate between samples containing **recombinant DNA**~~transgenic~~ protein above or below the ~~threshold~~**target level**, both a negative reference and a ~~threshold~~ reference containing a known proportion of **recombinant DNA**~~transgenic~~ grain should be assayed concurrently. The negative reference is a sample of the test matrix known to contain none of the protein analyte and is assayed to demonstrate that the method can distinguish between zero and the ~~threshold~~ **target** level. A sufficient number of these samples are run to ensure that assay sensitivity is adequate to determine whether the level in the test sample is greater or less than the ~~threshold~~ **target** level. During routine testing of bulk commodity samples, the lateral flow devices would typically be used without running the concurrent negative and ~~threshold~~ **target** reference samples.

VALIDATION OF A QUALITATIVE (THRESHOLD) PROTEIN-BASED METHOD

The same principles apply to qualitative protein-based testing as to qualitative PCR testing. These approaches, including calculation of false positive and false negative rates, can therefore be applied to protein-based methods. In general, due to the more reliable nature of protein-based lateral flow strip methods, they are not performed in duplicate on each sample. However, if ~~threshold~~ **target** ELISA testing is performed, duplicate wells should be used.

The same types of control samples, and criteria for acceptance/rejection of the result can be used as for qualitative PCR methods. The LOD is expressed as the amount of analyte at which the analytical method

detects the presence of the analyte at least 95% of the time (<5% false negative results). However, lateral flow strip tests are generally applied at test concentrations that are at least two fold (or more) above the LOD.

REFERENCES FOR ANNEX V

1. Guidelines for the Validation and Use of Immunoassays for Determination of Introduced Proteins in Biotechnology Enhanced Crops and Derived Food Ingredients. Lipton et al., Food and Agricultural Immunology, 2000, 12, 153-164.

2. Horwitz W; Protocol for the design, conduct and interpretation of method-performance studies. Pure and Applied Chemistry, 67, 331 (1995).

3. Residue Chemistry Test Guidelines OPPTS 860.1340 "Residue Analytical Method" United States Environmental Protection Agency, August 1996, (Mihaliak & Berberich, 1995).