

***United States – Measures Concerning the Importation, Marketing  
and Sale of Tuna and Tuna Products:***

***Recourse to Article 21.5 of the DSU by the United States  
Second Recourse to Article 21.5 of the DSU by Mexico***

**(DS381)**

Executive Summary of  
the United States of America

March 3, 2017

**TABLE OF ACRONYMS**

<b>Acronym</b>	<b>Full Name</b>
2016 IFR	Enhanced Document Requirements and Captain Training Requirements To Support Use of the Dolphin Safe Label on Tuna Products, 81 Fed. Reg. 15,444 (Mar. 23, 2016)
AIDCP	Agreement on the International Dolphin Conservation Program
DPCIA	Dolphin Protection Consumer Information Act
DSB	Dispute Settlement Body
DSU	Understanding on Rules and Procedures Governing the Settlement of Disputes
GATT 1994	General Agreement on Tariffs and Trade 1994
ETP	Eastern Tropical Pacific Ocean
FAO	United Nations Food and Agriculture Organization
FCO or Form 370	NOAA Fisheries Certificate of Origin
NMFS	National Marine Fisheries Service
NOAA	National Oceanic and Atmospheric Administration
RFMO	Regional Fishery Management Organization
TBT Agreement	Agreement on Technical Barriers to Trade
TTF	AIDCP tuna tracking form
WTO	World Trade Organization

## **I. THE AMENDED MEASURE IS CONSISTENT WITH ARTICLE 2.1 OF THE TBT AGREEMENT**

### **A. The Determination Provisions Are Even-Handed in Both Design and Application**

1. The Appellate Body found that the detrimental impact of the U.S. dolphin safe labeling measure did not stem exclusively from legitimate regulatory distinctions based only on the design of the determination provisions. As such, following the release of the Appellate Body report, the United States carefully reviewed the determination provisions, both in design and application.

2. As to design, the United States accepted the DSB recommendations and rulings and amended the determination provisions pursuant to the 2016 IFR in accordance with the Appellate Body's findings. Mexico does not appear to disagree with this conclusion, and thus the issue does not appear to be in dispute in this proceeding.

3. As to application, the United States also reviewed the readily available evidence to determine whether there is a basis for making a determination with respect to any particular fishery. With regard to the first prong of the determination provisions, the result of NOAA's analysis was that there is no evidence that there are fisheries that meet the first prong of the determination provisions, *i.e.*, in which there is a "regular and significant association between dolphins and tuna" similar to that in the ETP. With respect to the second prong, the available evidence suggests that certain gillnet fisheries in the Indian Ocean region meet that standard.

4. The text of the DPCIA is not explicit as to the metric whereby "regular and significant" dolphin mortality or serious injury should be assessed or as to the benchmark against which levels of dolphin mortality should be measured to determine whether they are "regular and significant." Consequently, it was necessary to consider what metric and benchmark were most in keeping with the purposes of the U.S. dolphin safe labeling measure, in light of the available evidence. Below, the United States explains: (1) the metric determined to be most appropriate; (2) the benchmark determined to be most appropriate; and, (3) the evaluation of different tuna fisheries on the basis of the available evidence.

5. First, in considering the most appropriate metric, NOAA concluded that a per set measure of dolphin mortality on a fishery-by-fishery basis reflects the frequency with which captains would have to make a determination that a dolphin was killed or seriously injured in a particular fishery. Such a metric is consistent with the purpose and structure of the U.S. measure because it assesses the effect of a tuna fishery on individual dolphins and because it is tailored to the frequency with which a vessel captain in a particular fishery would have to detect that a dolphin had been killed or seriously injured in a particular set. Specifically, where a captain would have to identify a dolphin mortality or serious injury more frequently, because more sets cause a direct dolphin harm, the determination provisions, if based on a per set metric, would provide that an observer certification (and enhanced tracking and verification) may be necessary for tuna product to meet the "dolphin safe" standard. Such a metric is also practical, as it used by many different regulating authorities, including by RFMOs, to assess the effect on dolphins (and other bycatch

species) of tuna fishing in a particular fishery, meaning that there is a considerable amount of per set data for different fisheries in different oceans that is readily available.

6. Second, having identified an appropriate metric, it was necessary to determine the appropriate benchmark against which fisheries would be evaluated to determine whether dolphin mortalities, on a per set basis, were “regular and significant.” In this regard, the United States recalled the suggestion of the Appellate Body that, to ensure even-handed treatment of different fisheries, the benchmark should refer to the ETP large purse seine fishery. NOAA determined that the most appropriate benchmark was a 20-year average of direct dolphin mortalities caused by dolphin sets in the ETP, beginning in 1997 and ending at the present day. In terms of promoting the objective of the measure while ensuring even-handed treatment of different fisheries, this approach has several advantages, namely: (1) averages are generally a more reliable basis on which to make scientific determinations than single-year figures; (2) it takes into account both the levels of mortality that were occurring at the time the enhanced observer and tracking and verification requirements were imposed and current levels; and, (3) it is conservative in nature, which is consistent with the objective of dolphin protection, because it takes into account any declines in direct mortalities due to dolphin sets in the ETP that have occurred over the past 20 years.

7. Third, on this basis, the United States considered the available fishery-specific evidence concerning per set mortalities in fisheries other than the ETP large purse seine fishery. As is shown by the relevant evidence on the record in this dispute, no evidence suggests that, on a per set basis, any other fishery causes close to the level of dolphin mortalities caused by dolphin sets in the ETP, as an average since 1997. Thus, no evidence suggested that any fishery for which fishery-specific evidence was available exhibited “regular and significant” dolphin mortality.

8. However, evidence from certain gillnet fisheries in the Indian Ocean area suggested that levels of mortality are occurring such that, if per set data were available, the per set mortality rate likely would meet or exceed the “regular and significant” standard. In particular, several exhibits presented in the first compliance proceeding suggested an alarming level of dolphin mortality was occurring in the Indian and Pakistani gillnet fisheries in the Indian Ocean, as well as in neighboring fisheries of other countries. The United States attempted to find per set data on these fisheries, but none was available. Consequently, the United States considered whether any alternative metrics might act as a proxy for per set data and enable an evaluation of those fisheries. NOAA determined that data were available to support evaluation under a dolphin bycatch rate metric, *i.e.*, the number of dolphins killed per ton of target catch (tuna) landed. While this is not a perfect metric for this analysis, it served as a reasonable proxy for the per set data. On September 28, 2016, NOAA issued a determination, on the basis of the best information available, that a “regular and significant” mortality of dolphins was occurring in these fisheries. The determination provided that any tuna product produced from these fisheries to be marketed as dolphin safe in the United States would have to be accompanied by a certification by an observer from a qualified and authorized observer program and a certification attesting to the catch documentation, the substance of the dolphin safe labeling standards, and the chain of custody information.

9. Thus, in the context of amending the design determination provisions, the United States evaluated their application based on an appropriate metric and benchmark. The evidence confirmed that, for all the fisheries for which per set mortality data is available, a positive determination was not required. For certain gillnet fisheries, however, this data was not available, but other relevant data suggested that these fisheries would meet the standard of “regular and significant mortality.” On this basis, and in the absence of contradictory information submitted by the countries, NOAA designated these fisheries. Thus, the application of the determination provision, like its design, is in compliance with Article 2.1.

10. In the first compliance proceeding, the design of the determination provisions was the sole basis on which the Appellate Body found that the detrimental impact of the measure did not stem exclusively from legitimate regulatory distinctions. In addition to addressing the DSB recommendation and the determination provisions, the 2016 IFR has also addressed the other concerns identified during the first compliance proceeding, even though these concerns did not form the basis for the DSB recommendation at issue. Looked at independently, the regulatory distinctions made by the eligibility criteria, the certification requirements, and the tracking and verification requirements are even-handed.

## **B. The Appropriate Legal Test for Even-Handedness in This Dispute**

11. For purposes of this dispute, as the Appellate Body has explained, the test for even-handedness is whether the regulatory distinctions, when viewed as a whole, are calibrated to the differences in risks to dolphins of overall harms between setting on dolphins in the ETP large purse seine fishery, on the one hand, and other fishing methods in other oceans on the other.

12. Mexico rejects the Appellate Body’s analysis. In particular, Mexico rejects the Appellate Body’s conclusions that: “there is a *special relevance* in these Article 21.5 proceedings in conducting an assessment” of whether the differences in labelling conditions are calibrated to the risks to dolphins; and a measure “*will not violate Article 2.1* if it is properly ‘calibrated’ to the risks to dolphins arising from different fishing methods in different areas of the oceans.”

13. First, Mexico gives no “special relevance” to the Appellate Body’s calibration test. Rather, in Mexico’s view, whether the differences in labeling conditions are calibrated to differences in overall harm to dolphins is merely “one element” of the legal test for even-handedness that the Panels should apply. Indeed, whether the measure is calibrated to the overall harm is not even the most important test under Mexico’s approach, as Mexico suggests that the key inquiry is whether differences in labeling conditions can be reconciled with the objectives of the measure (irrespective of the risk). Moreover, Mexico appears to reject the application of the Appellate Body’s calibration test for purposes of examining the certification requirements and tracking and verification requirements.

14. Second, Mexico implicitly – but repeatedly – alleges that the measure *will* violate Article 2.1 *even if* it is calibrated by contending that the measure is inconsistent with Article 2.1 for reasons unrelated to whether it is calibrated to risks of harm to dolphins. Thus, in its first written submission, Mexico argues that all three regulatory distinctions are inconsistent with Article 2.1 because they are “at odds with the objectives and the design, architecture and revealing structure

of the tuna measure.” While in its second submission Mexico claims that the measure is inconsistent with Article 2.1 because it is “inconsistent with the objectives of sustainable development,” as well as the fact that the measure does not apply “strict” certification and tracking and verification requirements to tuna product produced from fishing methods that result in above *de minimis* harms to dolphins or in fisheries subject to (allegedly) substandard municipal regulations.

### **C. Mexico’s Proposed Legal Tests for Even-Handedness Are Incorrect**

15. First, Mexico conceives of the even-handedness test as a multi-factor analysis where “one of the[] questions” is whether the measure is calibrated, and “[a]nother question is whether the regulatory distinctions are reconciled with the objectives of the measure.” In Mexico’s view, these two questions “do not create independent or discrete legal tests,” but rather are different elements that are weighed and balanced against one another to determine whether the measure is even-handed. Mexico misunderstands the Appellate Body’s even-handedness analysis as it applies in this dispute.

16. While the Appellate Body has recognized that there may be different approaches a panel could employ to test whether the challenged measure is even-handed, the facts and circumstances of a particular dispute will dictate the approach that is needed. *In this dispute*, the Appellate Body “considered appropriate an analysis” of whether the measure “is properly ‘calibrated’ to the risks to dolphins arising from different fishing methods in different areas of the oceans.” The fact that, in theory, there are different ways to test for even-handedness does not mean, as Mexico suggests, that the Panels must explore all these options in this dispute or that doing so would be appropriate.

17. Indeed, what Mexico seems to be arguing is that the Appellate Body has required the Panels to face the same situation that the first compliance panel did with regard to the certification requirements where that panel balanced the two analyses. The Appellate Body, however, *reversed* this analysis, finding that the majority did not conduct the appropriate analysis to determine the even-handedness of the certification requirements. Further, if Mexico were correct that the Appellate Body has required these Panels to engage in this multifactor balancing test, the Appellate Body’s analysis of the eligibility criteria would also have been fundamentally different.

18. In this dispute, the question of whether the detrimental impact stems exclusively from legitimate regulatory distinctions is not answered by conducting an assessment that fails to take into account differences in risks to dolphins. Rather, the question of even-handedness is answered by determining whether those differences in the regulatory distinctions “can be explained as being properly tailored to, or commensurate with, the differences in such risks in the light of the objective of protecting dolphins from adverse effects arising in different fisheries.” In this regard, it is clear that the calibration analysis, as set out by the Appellate Body, already takes into account the dual objectives of the measure. Specifically, the Appellate Body’s test examines whether the measure is calibrated to the risk of harm to dolphins as a way of explaining the differences in regulatory distinctions. Dolphin protection, including through consumer information, is the main substantive objective of the measure, and label accuracy is a

means of ensuring that end. Consequently, if the measure is appropriately calibrated to the risks to dolphins in different ocean areas, then the regulatory distinctions of the measure are consistent with the measure's objectives. Additionally, the fact that the Appellate Body analysis requires the same calibration test to determine whether the measure reflects arbitrary or unjustifiable discrimination for purposes of the chapeau of Article XX further confirms that, for purposes of this dispute, the analysis is not what Mexico claims.

19. Second, Mexico argues that the Panels should adopt different tests – not related to risks to dolphins, namely, “the risks of inaccurate information being provided” – to determine the even-handedness of the certification requirements and tracking and verification requirements than it should apply for the eligibility criteria. Mexico's arguments lack merit for several reasons. As explained above, the Appellate Body's calibration analysis already takes into account the objectives of the measure. Further, different tests do not apply to different regulatory distinctions. In the first compliance proceeding, the Appellate Body faulted that panel for applying a modified calibration test to the eligibility criteria and a different test to the certification requirements and tracking and verification requirements, emphasizing that the same test must be applied to each of these “cumulative and highly interrelated” regulatory distinctions. Mexico's approach urges the Panels to conduct an incorrect “segmented analysis” that the Appellate Body has already rejected. Moreover, the tests that Mexico claims must be used to determine whether the certification requirements and tracking and verification requirements are even-handed are incorrect. Specifically, Mexico's argument (that tuna product produced by a fishing method that causes above *de minimis* harm to dolphins must be ineligible for the label) is a rejection of the calibration analysis. Similarly, Mexico's other test, which posits that “strict” certification and tracking and verification requirements need to be applied to tuna caught in all fisheries where the level of applicable municipal fishery regulations falls below some (unspecified) minimum standard, is also a rejection of the calibration argument as it does not depend on an analysis of the differing risks to dolphins.

20. Third, Mexico's argument with regard to “environmental sustainability” is bereft of support, as to both the law and the facts.

#### **D. The Eligibility Criteria Are Calibrated to the Risk to Dolphins Posed by Different Fishing Methods**

21. The United States has established that the eligibility criteria are calibrated to the differences in risk to dolphins of overall harm. Specifically, the United States has established that: (1) setting on dolphins is a unique fishing method that is inherently unsafe for dolphins; (2) fishing methods that can produce tuna product eligible for the U.S. dolphin safe label do not pose equivalent risks to dolphins; and (3) the eligibility criteria are commensurate with the differences in risk.

##### **1. Setting on Dolphins Is a Unique Fishing Method that Is Inherently Unsafe for Dolphins**

22. The United States has shown that setting on dolphins is a fishing method that is particularly harmful to dolphins for three reasons: (1) it intentionally targets dolphins, such that

dolphins must be put at risk of direct and indirect harm in every dolphin set; (2) it causes a unique category of unobservable harms that may occur in every dolphin set, regardless of whether a dolphin is directly killed or injured; and, (3) it continues to cause a high level of direct dolphin mortalities.

23. First, it is well-established that setting on dolphins is the only fishing method that intentionally targets dolphins to catch fish. Thus, every dolphin set involves a sustained interaction between a large purse seine vessel (and its speedboats and often helicopters) and a herd of dolphins whereby the vessel chases about 600 dolphins for up to 2 hours, ultimately capturing about 300-400 of them. Other fishing methods, by contrast, can be used with no effect on dolphins at all. Furthermore, vessels may interact with a small number of dolphins only incidentally and those other fishing methods can be conducted without putting any dolphin directly in danger.

24. In this regard, Mexico is wrong to argue that the intentional nature of dolphin interactions is not relevant to the risk profile of setting on dolphins. Previous Appellate Body reports confirm that the analysis must be based on an assessment of the “overall levels of risk” caused by different fishing methods in different fisheries. The first compliance panel confirmed the relevance of the intentional nature of the dolphin interactions in dolphin sets, explaining that every set must involve a sustained interaction with hundreds of dolphins for up to several hours and that these interactions are inherently dangerous, as they can cause significant unobservable harms, as well as direct mortalities and serious injuries. There is, thus, no basis for excluding this feature of setting on dolphins from the analysis of the fishing method’s risk profile.

25. By contrast, as the first compliance panel recognized, other fishing methods in other fisheries do not intentionally target dolphins. In particular, as the United States has explained, other fishing methods can be used without harm to dolphins. Indeed, some fisheries – including certain handline, gillnet, longline, pole and line, and purse seine fisheries – pose no known, or only a remote, risk to any dolphins, due to the distribution of dolphins and the area where the fishery operates. In many other tuna fisheries, including in particular purse seine and longline fisheries, there is some known risk, but the vast majority of all sets occur without any dolphin interaction and, therefore, without putting any dolphin in danger. Thus, the vast majority of fishing activities in these fisheries involve no dolphins at all and, therefore, pose little or no risk of direct or indirect dolphin harms.

26. Second, setting on dolphins causes a unique category of indirect, unobservable harms due to the chase and encirclement process. Mexico argues that such effects are “speculative and unproven” and claims that, to the extent conclusions are drawn concerning indirect effects caused by dolphin sets, the same conclusions “must be presumed in relation to other fisheries and fishing methods.” Further, Mexico claims that the actions that cause the unique unobservable effects of dolphin sets are actually positive because they make the method sustainable. None of Mexico’s arguments have merit, and Mexico is wrong to try to “appeal” the DSB recommendations and rulings in these proceedings.

27. In this regard, Mexico is incorrect that if setting on dolphins is found to cause unobservable effects, then there must be a presumption that other fishing methods cause the



same effects. As the United States has explained, a significant body of peer-reviewed scientific literature concludes that setting on dolphins causes indirect, unobservable harms to dolphins due to the chase and encirclement process. Such harms include reproductive effects, calf-cow separation, and physical harms induced by stress. A 2007 article on the subject explained that a review of the existing literature showed, *inter alia*, that: (1) dolphin sets “entail well-recognized stressors in other mammals, especially wild animals,” and that typical responses “to such disturbances include changes in metabolism, growth, reproduction, and immune status, any of which, alone or in combination, could significantly affect survival and reproduction”; (2) samples from dolphins caught in dolphin sets “showed cell damage similar to that in heart muscle, indicative of a degree of capture myopathy that could lead to unobserved mortality in some cases”; and, (3) “developmental issues indicate that smaller calves (less than 1 year postpartum) may have more difficulty remaining associated with the mother during fishery activities.” A 2010 study explained that there is evidence that the ETP large purse seine fishery “has been a significant factor in the lives of dolphins since its inception” and, in particular, “is influencing reproduction in dolphin populations.” There are no such studies indicating similar indirect, unobservable mortalities caused by any other fishing method in any other fishery, and Mexico presents none.

28. Mexico has presented no evidence undermining previous findings that setting on dolphins causes a unique category of unobservable harms due to the chase itself that are not caused by other fishing methods. As a consequence, it is simply not possible for dolphin sets to be certified “dolphin safe” in the sense of having caused no harm to dolphins, since unobservable harms may occur in each set, even without any direct dolphin mortality, but cannot be seen by a captain or observer. Mexico’s attempts to portray dolphin sets as a sustainable, positively regarded fishing method are misleading and not relevant to whether the method is dolphin safe.

29. Third, setting on dolphins is a uniquely dangerous fishing method in terms of direct dolphin mortalities. Over the past decade, dolphin sets by ETP large purse seine vessels have caused several hundreds, and sometimes thousands, of direct dolphin mortalities per year. Controlling for the level of effort, in order to facilitate comparison across fisheries, these figures translate to between 69.4 and 126.3 dolphin mortalities per 1,000 dolphin sets. This level of dolphin mortalities, being caused by 80-90 vessels in approximately 10,000 sets per year, is generally unparalleled in other tuna fisheries.

## **2. Fishing Methods that Can Produce Tuna Product Eligible for the U.S. Dolphin Safe Label Do Not Pose Equivalent Risks to Dolphins**

30. Fishing methods that produce tuna product potentially eligible for the dolphin safe label generally pose a lower level of risk to dolphins than setting on dolphins in the ETP because: (1) they are not intrinsically harmful to dolphins and can be carried out without involving any dolphins; (2) they do not cause the types of unobservable harms caused by the chase and encirclement process regardless of whether dolphins are directly killed; and, (3) the levels of any direct dolphin mortality they may cause are generally lower, on a per set basis, than those caused by dolphins sets under the AIDCP and certainly are not high enough to counterbalance the unique risks posed by dolphin sets and thus equalize the risk profiles of dolphin sets and other fishing methods.

31. In this regard, Mexico is wrong to argue that “where there is credible evidence that dolphins have been harmed by a fishing method, it must be assumed that there are widespread direct and indirect harms *unless proven otherwise with absolute certainty.*” The Appellate Body found that calibration, not Mexico’s “zero tolerance” benchmark, is the applicable test for whether the U.S. measure, including the eligibility criteria, is even-handed. Mexico has put forward no evidence suggesting any other fishing methods are intrinsically dangerous to dolphins or cause such unobservable harms. Further, Mexico’s evidence does not show that, as a general matter, other fishing methods cause the level of direct dolphin mortalities caused by dolphin sets under the AIDCP.

32. With respect to purse seine sets other than dolphin sets, all the available set-by-set data confirms that such a fishing method causes a much lower level of direct dolphin mortality than dolphin sets. This is true for the Atlantic, Indian, western and central Pacific, and ETP Ocean purse seine fisheries where all of the evidence shows that levels of direct dolphin mortalities in these fisheries are well below those due to dolphin sets in the ETP. In this regard, the ETP large purse seine fishery itself provides the clearest example of why Mexico’s arguments fail. The evidence establishes that free school and floating object sets have accounted for over half of all sets in the ETP large purse seine fishery in the past decade but have caused only 0.2% of dolphin mortalities in the fishery – the other 99.8% being caused by dolphin sets. Mexico’s failure to respond to this evidence – or even acknowledge its existence at all – is telling. Purse seine sets other than dolphin sets thus cause significantly lower levels of direct mortality than dolphin sets and also put dolphins at risk less often and do not cause the unobservable harms that are caused by the chase and encirclement process.

33. With respect to longline fishing, the evidence establishes that this fishing method can be used without causing any harms to dolphins, and that the risks to dolphins from longlining in general in different areas of the oceans are significantly lower than the risks to dolphins from setting on dolphins in the ETP large purse seine fishery. While Mexico criticizes the United States for applying a “presumption,” Mexico fails to introduce evidence suggesting that the U.S. evidence on the record is not correct and representative. The United States has shown that, for example, in every tuna longline fishery for which data is available, the vast majority of sets (over 95 percent) occur without interacting with any dolphins. Further, the United States has shown that for every tuna longline fishery for which evidence is available, observed direct dolphin mortalities constitute small fractions of those caused by dolphin sets in the ETP on a fishery-by-fishery basis. Mexico has put forward no evidence contradicting the U.S. evidence.

34. With respect to pole and line fishing, Mexico does not even appear to contest that such a fishing method causes less overall harm to dolphins than does setting on dolphins in the ETP large purse seine fishery.

35. With respect to gillnet fishing, such a fishing method can produce tuna product that can be certified dolphin safe in the way that setting on dolphins cannot. Unlike in dolphin sets, dolphins are not an essential component of gillnet fishing and, therefore, gillnet sets, and even entire gillnet fisheries, can be conducted without interacting with and harming dolphins. Further, as the first compliance panel correctly found and the Appellate Body confirmed, gillnet fishing is not capable of causing the types of unobservable harms to dolphins that setting on dolphins

causes as a result of the “chase itself” even if no dolphins were directly observed to have been killed.

36. Mexico is further incorrect to argue that the levels of mortality caused by gillnets in the Indian Ocean gillnet fisheries supports a finding that all tuna product produced from gillnet fishing should be ineligible for the dolphin-safe label. The evidence on the record concerning the Indian Ocean gillnet fisheries is not suggestive of bycatch rates in all gillnet fisheries around the world. Dolphins are not evenly distributed throughout the world’s oceans, and different fisheries of the same gear-type can have vastly different bycatch levels depending on their area of operation and spatial overlap with dolphin populations. Some gillnet fisheries, in particular, are carried out in areas such that there is little or no known risk to any dolphin species. Further, there are techniques that can reduce dolphin interactions in gillnet fisheries, and thereby reduce the potential for dolphin harm. Thus, gillnet fishing overall presents a lower risk to dolphins than dolphin sets.

37. With respect to trawl fishing, the evidence establishes that trawl fishing does not cause a higher level of dolphin mortalities than dolphin sets in the ETP. In particular, studies of trawl fisheries indicate that, where they are used to catch tuna, bycatch is generally rare. One study commented that trawlers cause less mortality of marine mammals than other fishing methods, speculating that this might be due to “the disturbance caused by the trawling action at the bottom and at midwater warning cetaceans before they get caught.” With respect to tuna pelagic pair trawling in particular, the FAO explained: “In most cases [if it is] a single species fishery, bycatch rates of other species are low. . . . On few fishing ground[s], the incidental catch of dolphins and marine mammals creates some problems.”

38. With respect to handline fishing, Mexico has not even argued that such a fishing method causes a higher level of dolphin mortalities than setting on dolphins does in the ETP large purse seine fishery. Indeed, Mexico points to no dolphin mortalities or other harms reportedly caused by handline fishing. Rather, Mexico argues that handline vessels have been known to “chas[e]” dolphins and, therefore, if chasing dolphins is intrinsically harmful, “the tuna measure must disqualify tuna caught by handlines in association with dolphins in order to be even-handed.” However, none of Mexico’s evidence suggests that handline fishing in general or in the Indian Ocean is harmful to dolphins at all, let alone as harmful as setting on dolphins in the ETP large purse seine fishery.

### **3. The Eligibility Criteria Are Commensurate with the Differences in Risk**

39. The eligibility criteria are commensurate with the differences in risk to dolphins posed by setting on dolphins, on the one hand, and other fishing methods that produce tuna potentially eligible for the label, on the other. First, the eligibility criteria distinguish between the only fishing method that intentionally targets dolphins and those that do not. Because it intentionally targets dolphins, setting on dolphins is, by its very nature, inherently unsafe to dolphins. Other fishing methods by contrast, are not intrinsically dangerous, in that the intention of the fishing vessels is not to interact with dolphins (and, indeed, most sets occur without putting even one dolphin in danger). Second, the eligibility criteria are commensurate with the differences in risk

to dolphins of setting on dolphins and other fishing methods, as reflected in the number of dolphins directly endangered when such methods are employed. Setting on dolphins endangers, on average, hundreds of dolphins each and every time the method is employed. Other fishing methods, by contrast, only very rarely endanger a single dolphin, illustrated by the fact that vessels interact with dolphins in less than 1 percent of sets in nearly every fishery for which evidence is available. Third, the eligibility criteria are commensurate with the differences in risk posed by setting on dolphins and other fishing methods because they deny eligibility to a fishing method that may cause massive unobservable harms every time it is employed, irrespective of whether a dolphin has been killed or injured or whether AIDCP restrictions are applicable, while allowing eligibility for those fishing methods that generally do not cause many of these types of harms at all, and do not cause any of these harms without a dolphin being killed or seriously injured. Fourth, the eligibility criteria are commensurate with the differences in risk because they deny eligibility to a fishing method that causes a higher rate of observed mortalities and serious injuries while allowing eligibility for those fishing methods that cause a lower rate of observed mortalities and serious injuries.

40. For the reasons stated above, the eligibility criteria cannot support a finding of less favorable treatment under Article 2.1.

#### **E. The Certification Requirements Are Even-Handed**

41. Under the U.S. measure, as amended, all tuna product sold in the U.S. market as dolphin safe must be accompanied by a captain certification attesting that (1) “no purse seine net or other fishing gear was intentionally deployed on or used to encircle dolphins during the fishing trip in which the tuna were caught”; and (2) no dolphin mortality or serious injury occurred “in the sets or other gear deployments in which the tuna were caught.” Additionally, to be eligible for the label, tuna product produced from the ETP large purse seine fishery must be accompanied by certifications from an AIDCP-approved observer, as has been the case since the original measure came into effect. Moreover, under the 2016 IFR, captains operating outside the ETP large purse seine fishery are now required to certify completion of the NMFS dolphin-safe captain’s training course, regarding issues pertinent to the dolphin safe certification.

42. The certification requirements meet the test of calibration. Specifically, and as discussed above, the ETP large purse seine fishery has a special risk profile that is different from the risk profiles of other fisheries. The differences in the certification requirements are commensurate with these different risk profiles. This is the case for at least two reasons.

43. First, the differences in the certification requirements are commensurate with the differences in risk because the task of verifying that tuna meets the eligibility criteria is so much more difficult in the ETP large purse seine fishery than it is in other fisheries. That is to say, it is appropriate to require two certifiers (one of whom has to meet certain minimum education standards and has undergone some training) where the conditions facing the certifier are very difficult and to require only one certifier (who need not meet minimum education standards but is required to have taken a training course) where the conditions are less difficult. The ETP large purse seine fishery is fundamentally different from other fisheries in terms of the number of dolphins put at risk of mortality or serious injury by interacting with the vessel, fishing gear, etc.

and the frequency with which that interaction is taking place. Further, the ETP large purse seine fishery differs substantially from other fisheries in how this interaction occurs. ETP large purse seine vessels, in coordination with speedboats and a helicopter, engage in lengthy chases of large dolphin herds, which usually last 20-40 minutes but can take over two hours, with the entire process lasting another one-to-two hours following the end of the chase. Such a complex scene – in varying weather and ocean conditions – can make it very difficult for even the captain and a single observer to see every dolphin interaction throughout the entire process. There is no evidence that this type of interaction is repeated elsewhere in the world. In other fisheries, by contrast, where interactions with dolphins are generally accidental and are of limited scope and duration, captains are capable of determining the fate of the few dolphins they may encounter.

44. Second, the differences in certification requirements are commensurate with the differences in risk among fisheries because any difference in the “margin of error” resulting from the different requirements has a rational connection to the difference in risk, as discussed by the minority panelist in the first compliance panel’s report. That is to say, even if the conditions facing the certifiers in the ETP large purse seine fishery and other fisheries were the same (which they are not), and a captain working outside the ETP large purse seine fishery were, therefore, a less “sensitive” mechanism than an AIDCP observer, the regulatory distinction is calibrated (and thus even-handed) in tolerating a higher “margin of error” for the certifier where the risks are lower and tolerating a lower “margin of error” where the risks are higher. The United States has already demonstrated – indeed, the first compliance panel agreed – that the probability of dolphin mortality or serious injury is greater in the ETP large purse seine fishery than outside it. Thus, it remains the case that, as the minority panelist put it, the different certification requirements “represent[] a fair response to the different risk profiles existing in different fisheries, as established by the evidence,” even without taking into account that the certification requirements have narrowed since the first compliance proceeding. Taking the changes made by the 2016 IFR into consideration, it is even clearer that the certification requirements reflect a “fair response” to different risk profiles among fisheries, as established by the evidence on the record.

45. Thus, the certification requirements are calibrated to the risk profiles of different fisheries and, as such, are even-handed and thus cannot support a finding of less favorable treatment.

#### **F. The Tracking and Verification Requirements Are Even-Handed**

46. The purpose of the tracking and verification requirements of the NOAA tracking and verification regime is to distinguish between tuna product that meets the dolphin safe standard and tuna product that does not. The NOAA regime is composed of the interlocking elements of recordkeeping and reporting, physical segregation, verification, and sanctions. The 2016 IFR established additional chain of custody recordkeeping requirements for U.S. processors or importers, as applicable. U.S. processors and importers must collect and retain records regarding each custodian of the tuna or tuna product throughout the complete chain of custody, including storage facilities, transshippers, processors, re-processors, and wholesalers/distributors. These records must be sufficient for NMFS to conduct a trace-back to verify that any tuna product certified to NMFS as dolphin-safe, in fact, meets the dolphin-safe labeling requirements. Moreover, the recordkeeping must be sufficient for NMFS to trace any non-dolphin safe tuna

loaded onto the vessel back to one or more storage wells or other storage locations for a particular fishing trip to prove that such non-dolphin safe tuna was kept physically separate from dolphin-safe tuna from catch through unloading.

### 1. The Regulatory Differences Between the Two Regimes Are Narrow

47. The first compliance panel’s even-handedness analysis was based on a comparison of the two different sets of tracking and verification requirements. The first compliance panel concluded that the two regimes differed as to “depth,” “accuracy,” and “degree of government oversight.”

48. Depth. The first compliance panel used the term “depth” to refer to the point to which tuna can be traced back. The AIDCP regime requires that processed tuna product be traceable back to the AIDCP TTF, not the set or the well as the first compliance panel suggested. Thus, all that is required to be disclosed at the time of an audit is that the tuna in question was harvested on a particular trip covered by the TTF number, that it was caught in one of the sets listed on the TTF, and that it was stored in one of the wells listed on the TTF.

49. Accordingly, there is no practical difference in the tracking and verification requirements between the AIDCP and NOAA regimes as to “depth.” Both regimes have the same objective – to distinguish between dolphin safe and non-dolphin safe tuna, by being able to track and verify that the two types of tuna have been kept physically separate from one another from the vessel through processing. To do so, both regimes require a separate set of information for dolphin safe or non-dolphin safe tuna product (*i.e.*, different TTF pages for the AIDCP regime and different Form 370s (or equivalent) and captain certifications for the NOAA regime), to which NMFS can trace back the tuna.

50. Accuracy. The first compliance panel used the term “accuracy” to refer to “the degree of confidence that a particular captain (or, where applicable, observer) statement properly describes the lot of tuna to which it is assigned.” In the panel’s view, “[the AIDCP] tuna tracking forms . . . accompany particular catches of tuna throughout the fishing and production process, from the point of catch right through to the point of retail” and, “accordingly the identity of a particular batch of tuna can, in principle, always be established.” In contrast, for the NOAA regime, the first compliance panel questioned whether and how the “particular certificates are kept with particular lots of tuna up until the tuna reaches the canning plant.” The 2016 IFR directly addressed the first compliance panel’s concern with regard to “accuracy.” U.S. processors and importers now must maintain recordkeeping sufficient to allow NMFS to verify the dolphin safe status of tuna product. Such records must pertain to each custodian of the tuna or tuna product throughout the chain of custody, including storage facilities, transshippers, processors, re-processors, and wholesalers/distributors. In other words, this new recordkeeping requirement establishes a concrete legal obligation that the documentation attesting to whether the tuna is dolphin safe does, in fact, stay with the tuna throughout the supply chain.

51. Accordingly, there is no practical difference in the tracking and verification requirements between the AIDCP and NOAA regimes as to “accuracy.” Given that U.S. processors and U.S. importers must maintain records as to the complete chain of custody sufficient for NMFS to do a

complete trace back of the tuna product that is the subject of the verification, the legal requirements in place for both the AIDCP and NOAA regimes mean that “the degree of confidence that a particular captain (or, where applicable, observer) statement properly describes the lot of tuna to which it is assigned” will be the same.

52. Degree of Government Oversight. The first compliance panel used the phrase “degree of government oversight” to refer to “the extent to which a national, regional, or international authority is involved in the tracking and verification process.” In the panel’s view, in the AIDCP regime, “information concerning every stage of the tuna catch and canning process is made available to national and regional authorities, which must be sent copies of tuna tracking forms and are thus able to verify at any stage of the catch and canning process whether a particular batch of tuna is dolphin-safe.” By contrast, in the NOAA regime, the panel’s view was that “the United States has, as it were, delegated responsibility for developing tracking and verification systems to the tuna industry itself, including canneries and importers, and has decided to involve itself only on a supervisory and *ad hoc* basis through the review of monthly reports and the conduct of audits and spot checks.” Under the current measure, NMFS receives dolphin safe certifications for all tuna product sold on the U.S. market as dolphin safe. Moreover, all U.S. processors and importers marketing dolphin safe tuna product must now retain records such that the complete chain of custody for the tuna product, and the tuna contained therein, can be established. Thus, with respect to “government oversight,” the 2016 IFR narrowed the differences between the NOAA and AIDCP regimes.

## **2. The Differences in Requirements Are Calibrated to the Differences in Risk**

53. Like the certification requirements, the tracking and verification requirements are calibrated, and thus even-handed, because it is appropriate to use a more “sensitive” mechanism where the risks of dolphin mortality and serious injury are high, and a less “sensitive” mechanism where the risks of dolphin mortality and serious injury are low, as discussed by the minority panelist with regard to the certification requirements. The fact that the “mechanism” here occurs subsequent to the catch of the tuna does not mean that the calibration argument is rendered irrelevant to this stage of the analysis, as the Appellate Body has confirmed. Thus, any differences between the two “mechanisms,” *i.e.*, the tracking and verification systems, are small, particularly in light of the significant differences in risk between the ETP large purse seine fishery and other fisheries. With respect to depth, both the AIDCP and NOAA regimes require that tuna product that is “dolphin safe” (for purposes of their respective regimes) be traceable back to the harvesting vessel and trip and to the group of wells that held dolphin safe tuna. With respect to accuracy, both regimes require chain of custody recordkeeping sufficient to enable national authorities to trace a particular lot of tuna from harvesting through processing. With respect to government oversight, both regimes enable a government authority to obtain documentation “concerning every stage of the tuna catch and canning process” and thus both can “go behind” the dolphin safe certifications to the same extent.

## **G. The Measure, as a Whole, Is Even-Handed**

54. Although the United States considers it useful to discuss the issues on a distinction-by-distinction basis in order to provide a thorough explanation of how the calibration analysis applies in these proceedings, such a “segmented” approach is not the analysis that the Appellate Body has explained is necessary. Rather, the Appellate Body has stated consistency with the TBT Agreement and GATT 1994 must be determined based on a “comprehensive” analysis that takes into account the “cumulative and highly interrelated” nature of the different distinctions and “reconcil[es]” any different intermediate conclusions that may be drawn as to particular regulatory distinctions. In particular, the Appellate Body has stated that the fact that one regulatory distinction may not be “balanced” in relation to the risk does not mean the measure as a whole should be found inconsistent – it is necessary to look at whether the measure “as a whole” is commensurate with the risk.

55. This type of broad analysis makes sense. It is not the role of a WTO panel to step into the shoes of a regulator and find a measure inconsistent based on a hyper-technical investigation of whether the measure responds to every small detail of different circumstances around the world. The WTO Agreement has never been interpreted as requiring a measure to satisfy such a standard. Rather, the Appellate Body has indicated that the analysis should be the same type of analysis as other past panels have done, namely whether the measure, as a whole, is justified in light of the evidence available. In this light, the Appellate Body’s characterization of the calibration test – in particular, whether the regulatory distinctions are “commensurate” with the risk – is not materially different from how the previous proceeding’s minority panelist understood it. The question is simply whether the regulatory distinctions drawn by the measure reflect a “fair response to the different risk profiles existing in different fisheries, as established by the evidence.”

56. In fact the measure – when viewed as a whole – reflects such a “fair response” to the evidence. In particular, the evidence establishes that the ETP large purse seine fishery has a higher risk profile than other fisheries because a uniquely dangerous fishing method is widely employed in that fishery that is not employed in other fisheries. The measure recognizes the unique risk to dolphins that the intentional chase and capture of dolphins poses and draws distinctions between fishing methods and fisheries in light of that difference in risk. The difference in risk “explains” the distinctions contained in the measure, and the measure is, in turn, calibrated to the risk. Accordingly, the measure is consistent with Article 2.1.

### **III. THE MEASURE IS CONSISTENT WITH ARTICLE XX OF THE GATT 1994**

57. For purposes of this dispute, the Appellate Body has stated that the question to be answered to determine whether the measure reflects “arbitrary or unjustifiable discrimination,” as understood in the chapeau of Article XX, is the same question that needs to be answered to determine whether the measure is even-handed under Article 2.1. Namely, is the measure calibrated to differences in risk of overall harm between setting on dolphins in the ETP large purse seine fishery and other fishing methods employed in other fisheries? As discussed, the measure is, in fact, so calibrated. As such, any inconsistency with GATT 1994 Articles I:1 and III:4 is justified under Article XX.