

Proposal to introduce a “Quality Score” in EXFOR

E. Dupont, A.J. Koning¹, N. Otsuka²

OECD NEA Data Bank, Issy-les-Moulineaux, France

¹ Nuclear Research and Consultancy Group, Petten, The Netherlands

² IAEA Nuclear Data Section, Vienna, Austria

1. Introduction

The WPEC Subgroup 30 (SG30) studied different ways to improve the accessibility and the quality of the EXFOR database [1]. During its lifetime (2007-2010), several batches of error reports have been sent to the NRDC [2-5]. The final SG30 report describes the translation process from EXFOR into more user-friendly tabulated data files, as well as the various methods developed to verify and correct the contents of the database. SG30 contributed to the long-term objective to establish EXFOR as an easily accessible and correct database. However, SG30 achievements were only the “end of the beginning” and SG30-like activities are still continuing at NRG and Data Centres, e.g.

- Distribution of EXFOR in C4 computational format (V. Zerkin, IAEA-NDS).
- Correction system to collect and apply evaluator’s feedback on C4 data (V. Zerkin, IAEA-NDS).
- Compilation and monitoring by IAEA-NDS of coding mistakes reported by users and NRDC staff [6].
- Various checking performed at the IAEA-NDS on incident energies coded in MeV instead of keV, level energies higher than 20 MeV or lower than 10 keV, reactions violating charge or mass conservation, data above threshold of second level excitation without specification of level energy.
- Implementation of SG30 methods at the NEA DB as described in references [7] and in appendix A [8].
- Automatic test of C4 data and comparison with TALYS/TENDL at NRG as described in references [1,5] and in appendix A [8].

Today, the NRG review of EXFOR data goes far beyond the initial scope of SG30 and already includes a quality assignment of the EXFOR data.

The purpose of this working document is to present this quality classification to NRDC members and to discuss the opportunity and the technical solutions to store this valuable information in the EXFOR database.

2. Quality assignment

In a first step, quality scores are assessed automatically, by comparing the experimental values of the subentry with model calculations or other data libraries. In a second step, paper reviews and possible corrections are performed, and subentries may be finally placed in a scoring class which is different from the initial one. The appendix B describes in more details a quality classification that puts this into effect, as well as a strategy to come to an improved, verified EXFOR library. In this summary, we only briefly describe this quality classification system.

2.1 Calculation of the deviation factor

Automated comparison between evaluated and experimental data is done using the three following deviation factors, which are available per data point, per data set (as below), per reaction, per projectile and for the whole library:

$$F = 10 \sqrt{\frac{1}{N} \sum_i^N \left(\log \frac{\sigma_T^i}{\sigma_E^i} \right)^2}, \quad \chi^2 = \frac{1}{N} \sum_i^N \left(\frac{\sigma_T^i - \sigma_E^i}{\Delta \sigma_E^i} \right)^2, \quad \Delta = \frac{1}{N} \sum_i^N |\sigma_T^i - \sigma_E^i|,$$

where T stands for TALYS/library and E for experimental data (with N points).

2.2 Quality classification

The following quality classes are defined:

- Class 1 ($1 \leq F \leq 1.2$) for data deviating in average by less than 20%;
- Class 2 ($1.2 < F \leq 2$) for data deviating in average between 20% and a factor 2;
- Class 3 ($F > 2$) for data deviating by more than a factor 2.

Class 3 data should always be reviewed by an expert, while class 2 data should be reviewed only if Δ is more than 10% of the non-elastic cross-section and $\chi^2 > 20$.

In addition to this quantitative classification, subentries are sorted in 4 main categories:

- “T” for automated comparison of data with TALYS/TENDL and other evaluated libraries;
- “R” for data reviewed against the original publication;
- “E” for data erroneously stored in EXFOR;
- “N” for data that could not be cross-checked with the original publication (yet).

2.3 Review of activation cross-sections

For now 1 year, informal collaboration has started between NEA DB, NRG and IAEA-NDS to review all 6827 (n,2n), (n,p), (n, α) partial and total cross-sections available in the EXFOR database, i.e. 2866 (n,2n), 2561 (n,p) and 1400 (n, α) subentries (cf. Table 1). All these subentries were assigned a quality score after cross-checking with information from other measurements, libraries and/or calculations, as well as verification of the associated publications.

Table 1: Total number of neutron-induced cross-section subentries available in XC4 format.

Reaction	All	Reaction	Number	Reaction	Number	Reaction	Number
Composite	5791	Total	4493	Elastic	1118	Non-elastic	425
(n, γ)	5699	(n, γ) (tot)	4932	(n, γ)g	287	(n, γ)m	480
(n,f)	1259						
(n,n')	579	(n,n') (tot)	303	(n,n')g	7	(n,n')m	269
(n,n' _k)	1162	(n,n' ₁)	517	(n,n' ₂)	160	(n,n' ₃)	64
(n,2n)	2866	(n,2n) (tot)	1677	(n,2n)g	402	(n,2n)m	787
(n,p)	2561	(n,p) (tot)	1878	(n,p)g	219	(n,p)m	464
(n, α)	1400	(n, α) (tot)	1081	(n, α)g	96	(n, α)m	223
part prod	515	(n,xn)	20	(n,xp)	94	(n,x α)	181
Other	665	(n,3n)	127	(n,n α)	83	(n,np)	116
Total	23490						

As an example, Figure 1 shows a comparison between experimental and evaluated data for the $^{72}\text{Ge}(n,p)^{72}\text{Ga}$ cross-section. This is judged together with the F-values given in Table 2. The data set measured by Paul, initially categorized in Class 2, was reviewed by checking the publication, and both the value and the reaction identifier stored in EXFOR were found to be correct. Finally, this data set was flagged as R2 (Reviewed). This exercise was extended to all (n,p), (n, α) and (n,2n) cross-sections.

Table 2: F-values for all $^{72}\text{Ge}(n,p)$ data sets.

Author	F	TALYS	ENDF	JENDL	JEFF	CENDL	EAF	TENDL
Casanova	1.08	1.13	1.09	1.09	1.02	1.01	1.02	1.13
Konno	1.11	1.05	1.08	1.08	1.12	1.16	1.18	1.04
Changlin Lan	1.13	1.20	1.13	1.13	1.08	1.05	1.06	1.20
Hoang	1.15	1.19	1.18	1.18	1.13	1.05	1.03	1.18
Vinitskaya	1.16	1.16	1.05	1.05	1.13	1.24	1.23	1.15
Qaim	1.16	1.14	1.06	1.06	1.14	1.24	1.23	1.14
Rieppo	1.20	1.18	1.28	1.28	1.19	1.09	1.10	1.19
Zhong-Sheng	1.26	1.18	1.22	1.22	1.30	1.34	1.36	1.17
Wood	1.35	1.38	1.42	1.42	1.32	1.24	1.24	1.39
Vlastou	1.44	1.05	1.64	1.64	1.62	1.25	1.51	1.04
Paul	1.85	1.88	1.96	1.96	1.82	1.69	1.71	1.89
Average	1.29	1.21	1.33	1.33	1.31	1.21	1.30	1.21

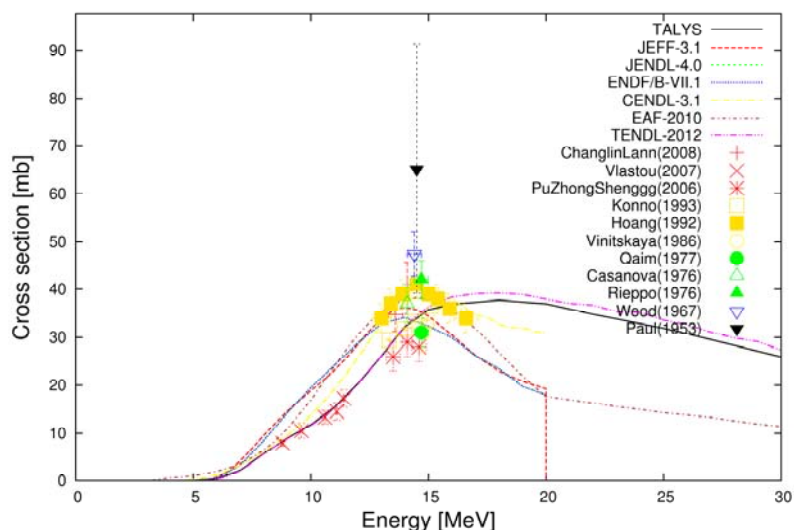


Figure 1: Comparison of experimental and evaluated data for $^{72}\text{Ge}(n,p)^{72}\text{Ga}$.

Table 3 summarizes the current quality scoring for all (n,2n), (n,p) and (n, α) cross-sections. After completion of the full review the T3 line will be basically empty and the N3 articles will have to be scanned and reviewed with the highest priority. Of course, all subentries in the category E will have to be corrected.

Table 3: Quality scoring of (n,2n), (n,p) and (n, α) as of April 22, 2013

Score\reaction	(n,2n)	(n,p)	(n, α)
T1	684	320	153
T2	402	452	270
T3	100	95	127
R1	309	265	84
R2	292	454	179
R3	87	162	143
N1	495	277	116
N2	380	386	190
N3	80	120	115
E1	0	0	0
E2	4	5	1
E3	6	11	2

3. Conclusion

As already seen in the course of Subgroup 30 activities, there is a number of data in the EXFOR database that differ significantly from other measurements or theoretical values (cf. line "R3" in Table 3). Although these data are correct and properly stored in the EXFOR database, it would be useful to inform the users that they differ significantly from other measurements or theoretical values. This additional information on the quality of the data set should be provided in an unambiguous and systematic way using a specific Quality keyword.

References

- [1] Nuclear Energy Agency (NEA), "Quality improvement of the EXFOR database", *International Evaluation Co-operation*, Vol. 30, NEA/WPEC-30, OECD/NEA, Paris (2011). See www.oecd-neo.org/science/wpec.
- [2] "NRDC action to the list of EXFOR outliers", NRDC working paper WP2008-3 (2008).
- [3] "EXFOR Outliers (Parts 4 and 5)", NRDC working paper WP2010-10 and memo CP-D/623 (2010).
- [4] "Automatic test of EXFOR with TALYS", NRDC working paper WP2010-11 and memo CP-D/627,633 (2010).
- [5] "Follow-up of WPEC Subgroup 30 activities on quality improvement of the EXFOR database", NRDC working paper WP2011-17 (2011).
- [6] www.nds.iaea.org/nrdc/error.
- [7] "Statistical methods for the verification of databases – Application to the international database of experimental nuclear reaction data (EXFOR)", NEA News Magazine 29.1, 30 (2011). See Publications at www.oecd-neo.org/nea-news.
- [8] O. Zeydina *et al.*, "Cross-checking of large evaluated and experimental databases", Int. Conf. on Nuclear Data for Science and Technology (ND2013), New York, USA, March 4-8, 2013. See Appendix A.

Cross-checking of large evaluated and experimental databases

O. Zeydina,¹ A.J. Koning,² N. Soppera,³ D. Raffanel,¹ M. Bossant,³ E. Dupont,^{3,*} and B. Beauzamy¹

¹*Societe de Calcul Mathematique, Paris, France*

²*Nuclear Research and Consultancy Group, Petten, The Netherlands*

³*OECD Nuclear Energy Agency Data Bank, Issy-les-Moulineaux, France*

(Dated: April 19, 2013)

This paper presents mathematical methods for the verification of large experimental and evaluated nuclear databases (e.g. EXFOR, JEFF, TENDL). These methods allow to assess the overall consistency of the data and to detect aberrant values in both evaluated and experimental database.

I. INTRODUCTION

The Nuclear Energy Agency (NEA) Data Bank is a member of the network of Nuclear Reaction Data Centre [1] that maintains large nuclear databases and makes them available to scientists and engineers worldwide. In the continuity of WPEC Subgroup 30 activities [2], the NEA Data Bank, the Societe de Calcul Mathematique (SCM) and the Nuclear Research and Consultancy Group (NRG) developed statistical methods to help improve the quality of these databases.

An outline of the NEA-SCM procedure to check the mutual consistency of experimental data is described in reference [3]. The focus of the present contribution is to describe methods developed by NEA-SCM and NRG to check the consistency of data from different evaluated libraries (e.g. JEFF, TENDL) and to cross-check these evaluated data with experimental data (EXFOR).

This approach aims to use the valuable information stored in both experimental and evaluated database. It helps assess the quality of evaluated data by comparison with experimental data (and vice-versa) and helps identify flaws in evaluated or experimental database.

II. DATABASE DESCRIPTION

In the present work, we only consider the case of differential cross-section $y = f(x)$, where x is the energy of the incident particle.

Experimental data (EXFOR) are the result of measurements and considered as a set of points, whereas the data stored in evaluated files are the result of modeling or expert evaluation and considered as continuous functions that can be complemented by additional points.

In addition, there are some particularities that must be taken into account: (i) an experimental point may be associated with an uncertainty and/or a resolution; (ii) both experimental and evaluated data may contain resonance regions characterized by large fluctuations in cross-section values; (iii) low value of cross-sections may lead to large relative discrepancies.

III. METHOD DEVELOPED BY NEA-SCM

The main idea of the consistency assessment is to calculate the mean distance between (i) two curves in the case of evaluated data; (ii) a curve and a set of points when comparing evaluated and experimental data.

The comparison is performed separately for each nuclear reaction by means of relative and absolute “distances” between experimental points and each of the evaluated curves. Subsequently, the available nuclear reactions are rated by combining the two types of distances. The same procedure is applied when checking the mutual consistency of evaluated data: each library is compared with all available evaluations of the same reaction and then rated according to its similarity with the others.

The absolute distance between two libraries describing the same reaction is calculated as the integral of the difference between two continuous functions:

$$d = \int_a^b |f(x) - g(x)| dx. \quad (1)$$

The absolute distance between a curve and a set of points is the average of the differences observed at the same abscissa (energy):

$$d = \frac{\sum_{i=1}^N |y_i - y'_i|}{N}, \quad (2)$$

where y_i is the experimental value, y'_i is the evaluated (interpolated) value, and N is the number of experimental

* Corresponding author: emmeric.dupont@oecd.org

data points. The relative distance is deduced by dividing the absolute distance by the average of compared values. This method is implemented in seven steps.

(1) *Representation of the data*: An essential point when representing the data is to define the optimal scales that allows plotting the data as evenly spaced as possible [3].

(2) *Energy range*: For a given reaction only energy values common to all data sets are taken into account.

(3) *Discretization*: The energy range is discretized with 50 subintervals of the same size.

(4) *Construction of the resonance indicator*: Before making a comparison, resonance regions must be distinguished. We calculate a “relative variance” for each discretization interval, i , using the formula:

$$var_i^{rel} = \frac{\sum_{n=1}^{N_i} \frac{|Y_n - \bar{Y}|}{(Y_n + \bar{Y})/2}}{N_i}, \quad i = 1, 2, \dots, \quad (3)$$

where N_i is the number of measurements; Y_n and \bar{Y} are respectively the cross-section and its average value. A resonance region is defined by a “relative variance” larger than a given threshold, which was empirically set to 0.18 (see example in Fig. 1).

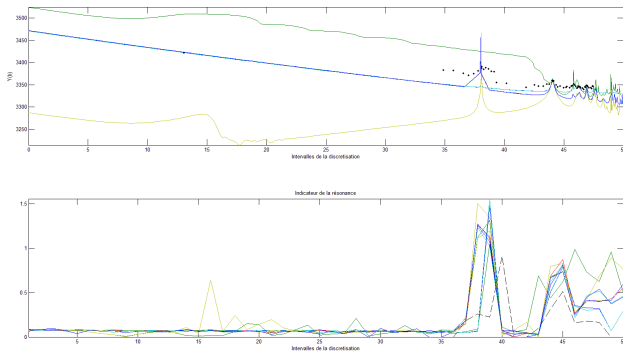


FIG. 1. Data available in the EXFOR database (points) and in evaluated libraries (lines) for the reaction $^{11}\text{B}(n, \gamma)$ (top) and the values of relative variance for each interval (bottom).

(5) *Consistency of evaluated and experimental data*: The comparison is performed subject to the value of the resonance indicator: (i) in a resonance region, the distance is computed as the difference of integrals:

$$d = \left| \int_a^b f(x) dx - \int_a^b g(x) dx \right|, \quad (4)$$

where the integral for EXFOR data is made using interpolation between points from the same data set; (ii) otherwise Eq. (2) is applied.

(6) *Mutual consistency of evaluated data*: (i) in a resonance region, the distance is computed using Eq. (4); (ii) otherwise Eq. (1) is used. Figures 2 and 3 present the distances calculated for the reaction given in Fig. 1. The 2nd and the 6th evaluated libraries are the worst cases compared to other libraries and experimental data (see also curves in Fig. 1).

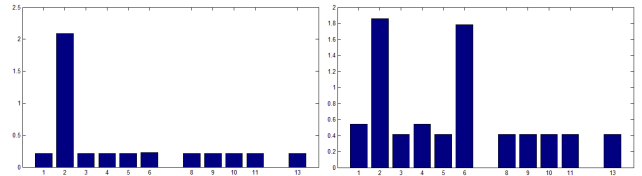


FIG. 2. Assessment of the self-consistency of evaluated data: absolute (left) and relative distances (right).

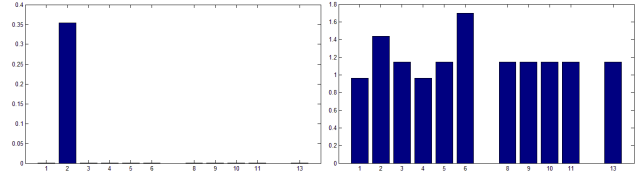


FIG. 3. Assessment of the consistency between measured and evaluated data: absolute (left) and relative distances (right).

(7) *Classification of nuclear reactions*: In order to rank nuclear data according to consistency criteria, an indicator combining the absolute and relative distances is implemented as follow: (i) using steps (5) and (6), compute the absolute (*abs*) and relative (*rel*) distances for each single library (average of all available libraries if needed); (ii) calculate $\log_{10} abs$; (iii) using all examined reactions, normalize these values between [0.005-1.005]; (iv) the final indicator is given by $rel \times (\log_{10} abs)^{norm}$. Note that this indicator is still relative in the sense that it only helps identifying on a consistency scale the best and the worst cases among all cases studied.

IV. METHOD DEVELOPED BY NRG FOR QUALITY ASSESSMENT AND CORRECTION

To efficiently verify and improve the EXFOR database, a quality score should be given to each subentry. This is no small task given the total size of the EXFOR database, which is summarized in Table I. Out of the 134941 subentries (data sets), 83084 have been translated by the IAEA into the extended computational format (XC4), which makes EXFOR more easily readable for further processing. Out of this, 97% has been identified as usable data and categorized in a directory-structured database developed at NRG, with the preliminary name “Newbase”. Finally, 52% of these subentries, at the moment cross sections only, have been compared with the latest version of

TABLE I. Total contents of EXFOR and derived databases.

Database	# entries	# subentries	# data points
EXFOR	19764	134941	11882567
XC4	13621 (69%)	83084 (62%)	7572483 (64%)
Newbase	13342 (98%)	80660 (97%)	7471028 (99%)
TALYS + lib.	7829 (59%)	41958 (52%)	4062317 (54%)

TALYS and/or the ENDFB/VII.1, JEFF-3.1.1, JENDL-4.0, CENDL-3.1, EAF-2010, IRDFF-1.0 and TENDL-2012 databases, in pointwise format. In other words, for 4062317 data points we have an automated comparison between experiment and evaluation. The main goodness-of-fit estimator we are using is the F-factor

$$F = 10 \sqrt{\frac{1}{N} \sum_i^N \left(\log \frac{\sigma_T^i}{\sigma_E^i} \right)^2}, \quad (5)$$

though we also monitor the χ^2 estimator,

$$\chi^2 = \frac{1}{N} \sum_i^N \left(\frac{\sigma_T^i - \sigma_E^i}{\Delta \sigma_E^i} \right)^2, \quad (6)$$

and the absolute deviation

$$\Delta = \frac{1}{N} \sum_i^N |\sigma_T^i - \sigma_E^i|. \quad (7)$$

In these equations, the subscript T stands for TALYS/library and E for experimental. In all cases, we average over the number of energy points, N, in each data set. The F-factor can thus be given per data point, in which case it is just a twisted C/E factor, per data set summed over all points, per reaction summed over all data sets and/or nuclides, per projectile summed over all reactions, and finally for the whole EXFOR database, for each of the above mentioned libraries.

It is quite clear that if we want a *quality score* for each data set in such a large database, that initially these scores will be assigned automatically. Next, paper reviews and possible corrections may be performed, and subentries may be placed in a scoring class which is different from the initial one. A preliminary quality classification is proposed which puts this into effect:

- Class 1: $1 \leq F \leq 1.2$
- Class 2: $1.2 < F \leq 2$
- Class 3: $F > 2$

At the moment, we feel that quality class “3” should always be reviewed by checking the publication for possible compilation mistakes, while class “2” should be reviewed if in addition D is more than 10% of the total nonelastic cross section and if χ^2 is larger than 20. After that, the data set gets a “R” (reviewed) flag or an “E” (error) flag.

As an example, Fig. 4 shows a comparison between experimental data and libraries for $^{72}\text{Ge}(n,p)$. This is judged together with the numerical values of Table II. On the basis of that the data set by Paul was reviewed, and in this case it was concluded that it was indeed the

reported value. This exercise could be performed for the entire EXFOR database.

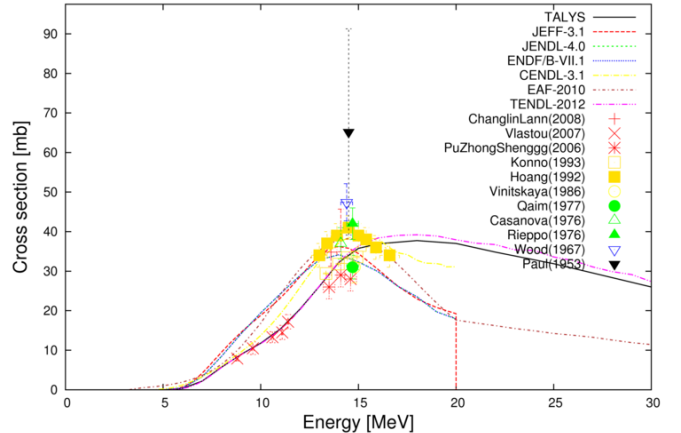


FIG. 4. Comparison of data for $^{72}\text{Ge}(n,p)$

TABLE II. F-values for all $^{72}\text{Ge}(n,p)$ data sets.

Author	F	TALYS	ENDF	JENDL	JEFF	CENDL	EAF	TENDL
Casanova	1.08	1.13	1.09	1.09	1.02	1.01	1.02	1.13
Konno	1.11	1.05	1.08	1.08	1.12	1.16	1.18	1.04
Changlin Lan	1.13	1.20	1.13	1.13	1.08	1.05	1.06	1.20
Hoang	1.15	1.19	1.18	1.18	1.13	1.05	1.03	1.18
Vinitkaya	1.16	1.16	1.05	1.05	1.13	1.24	1.23	1.15
Qaim	1.16	1.14	1.06	1.06	1.14	1.24	1.23	1.14
Rieppo	1.20	1.18	1.28	1.28	1.19	1.09	1.10	1.19
Zhong-Sheng	1.26	1.18	1.22	1.22	1.30	1.34	1.36	1.17
Wood	1.35	1.38	1.42	1.42	1.32	1.24	1.24	1.39
Vlastou	1.44	1.05	1.64	1.64	1.62	1.25	1.51	1.04
Paul	1.85	1.88	1.96	1.96	1.82	1.69	1.71	1.89
Average	1.29	1.21	1.33	1.33	1.31	1.21	1.30	1.21

V. CONCLUSIONS

The work presented here is a follow-up of previous activities initiated in the framework of WPEC Subgroup 30 [2]. The performance of the NEA-SCM method was verified on a selection of about two hundred cross-sections, which were all successfully analysed and ranked. Further testing is under way using full JANIS databases [4], including future JEFF test-libraries. The method developed at NRG is being used to assign quality scores to EXFOR data set with the ultimate objective to improve the accuracy of TENDL libraries. These initiatives will help improve the quality of the EXFOR database, JEFF and TENDL libraries for the benefit of all nuclear data users.

[1] N. Otuka *et al.*, NRDC, these proceedings.
 [2] A.J. Koning *et al.*, NEA/WPEC-30 (2011).

[3] E. Dupont *et al.*, NEA NEWS MAGAZINE **29.1**, 30 (2011).
 [4] N. Soppera *et al.*, JANIS, these proceedings.

Quality assignment of the EXFOR database

Version April 22 2013

Arjan Koning (NRG)

1 Quality scoring

To efficiently verify the EXFOR database, a quality score should be given to each subentry. Initially these scores will be given automatically, by comparing the experimental values of the subentry with model calculations or other data libraries. Next, paper reviews and possible corrections may be performed, and subentries may be placed in a scoring class which is different from the initial one. First, a quality classification is proposed which puts this into effect. Next, the order of steps to come to an improved, verified EXFOR library is outlined. In a later section, the numerical criteria for quality classes are defined.

1.1 Quality classes

Subentries which are not reviewed or not automatically compared (blank)

blank **Neither reviewed nor compared with models/evaluations.**

The subentry is not (yet) cross-checked with information from other measurements, libraries and/or calculations. This is the default quality score.

Subentries which are automatically compared with TALYS or data libraries (T)

T1 Automatically compared with models/libraries: small deviations.

The subentry contains (very) probably the reaction and data measured by the author, and although the associated publication has not been checked by the reviewer, the quantities have central values and uncertainties which are close to other measurements, libraries and/or calculations.

T2 Automatically compared with models/libraries: questionable deviations

The subentry contains maybe the reaction and data measured by the author, and the associated publication has not (yet) been checked by the reviewer. The quantities have central values and uncertainties which are deviating to some extent from other measurements, libraries and/or calculations.

T3 Automatically compared with models/libraries: strong deviations The subentry contains probably not the reaction and data measured by the author, and the associated publication has not (yet) been checked by the reviewer. The quantities have central values and uncertainties which are strongly deviating from other measurements, libraries and/or calculations.

Subentries which are reviewed by checking the publication (R or N)

R1 Paper reviewed: small deviations.

The subentry contains certainly the reaction and data measured by the author, since the associated publication has been checked by the reviewer. The quantities have central values and uncertainties, which are close to other measurements, libraries and/or calculations.

R2 Paper reviewed: questionable deviations

The subentry contains certainly the reaction and data measured by the author, since the associated publication has been checked by the reviewer. The quantities have central values and uncertainties which are deviating to some extent from other measurements, libraries and/or calculations.

R3 Paper reviewed: strong deviations

The subentry contains certainly the reaction and data measured by the author, since the associated publication has been checked by the reviewer. The quantities have central values and uncertainties which are strongly deviating from other measurements, libraries and/or calculations.

N1 Automatic score T1, but pdf of paper not available for checking

N2 Automatic score T2, but pdf of paper not available for checking

N3 Automatic score T3, but pdf of paper not available for checking

Subentries which contain errors and require specified action (E)

E1 Error: subentry contains other/related quantity.

The subentry does not contain the reaction or data measured by the author, but (very) probably a related quantity, since that related quantity agrees with other measurements, libraries and/or calculations and with the associated publication, which has been checked by the reviewer. Action: confirmation and correction by Data Centres.

E2 Error: subentry contains correct reaction, but wrong values.

The subentry contains (very) probably the reaction measured by the author, since the associated publication has been checked by the reviewer, but the values are wrong (e.g. wrong units). Action: confirmation and correction by Data Centres.

E3 Error: subentry contains data which are unrecognizable

The subentry contains reaction and data that do not agree at all with other measurements, libraries and/or calculations. Furthermore, no origin of the value or alternative meaning for the value could be found, although the associated publication has been checked by the reviewer. Action: further analysis by Data Centres.

For each subentry, a data block will be created including the basic information of the subentry, the date of the review and the quality score.

2 Various stages of quality assignment

Stage 1

All X4 entries that can be automatically compared with TALYS will get a quality score T1, T2 or T3. A quality score T1 will probably not be changed anymore. Only if unexpected new information comes to surface, from either the experimental or modelling side, this may change. Hence, the result of this stage is:

- T1: Definite assignment in database
- T2+T3: Should be reviewed.

Stage 2

The papers of subentries with quality score T2 and T3 will be reviewed. In the course of time, the boundary between T1, T2 and T3 may be altered. This depends on the number of false alarms in the T2 class, which determines whether the decision for a paper review should be more or less strict. After paper review, a subentry with an initial quality score of T2 will end up in R2, N2, E1 or E2, while a subentry with an initial quality score of T3 will end up in R3, N3, E1, E2 or E3. Hence, the result of this stage is:

- R2+R3: Definite assignment in database
- E1+E2+E3: Should be corrected
- N3 (and maybe N2): should have priority for acquiring the pdf file of the paper

Stage 3

Cases with quality score E1, E2, E3 will result in a message to the Data Centres with a recommended correction. After this correction, these subentries will be reviewed again after which they may be upgraded to R1, R2 or R3 in the next X4 update. Hence, the result of this stage is:

- R1+R2+R3: Definite assignment in database

Stage 4

The final quality scores, i.e. after all corrections, will be either 0, T1, N1, N2, N3, R1, R2 or R3.

Reaction	All	Reaction	Number	Reaction	Number	Reaction	Number
Composite	5791	Total	4493	Elastic	1118	Non-elastic	425
(n, γ)	5699	(n, γ) (tot)	4932	(n, γ)g	287	(n, γ)m	480
(n,f)	1259						
(n,n')	579	(n,n') (tot)	303	(n,n')g	7	(n,n')m	269
(n,n' _k)	1162	(n,n' ₁)	517	(n,n' ₂)	160	(n,n' ₃)	64
(n,2n)	2866	(n,2n) (tot)	1677	(n,2n)g	402	(n,2n)m	787
(n,p)	2561	(n,p) (tot)	1878	(n,p)g	219	(n,p)m	464
(n, α)	1400	(n, α) (tot)	1081	(n, α)g	96	(n, α)m	223
part prod	515	(n,xn)	20	(n,xp)	94	(n,x α)	181
Other	665	(n,3n)	127	(n,n α)	83	(n,np)	116
Total	23490						

Table 1: Total number of neutron-induced *cross section* subentries available in XC4 format.

3 Goodness-of-fit estimators for EXFOR

To judge the quality of an experimental data point, we can compare it with predictions from TALYS and, for neutrons, also with values as evaluated in the nuclear data libraries ENDFB/V-II.1, JENDL-4.0, JEFF-3.1.2, CENDL-3.1, IRDFF1.0, EAF-2010 and TENDL-2012. Moreover, one can compare with other measurements of the same reaction. The existing nuclear data libraries + TALYS should be able to give a reasonably good prediction of many reaction data, and obviously we will always try to extend such predictions to as many reactions as possible in future versions. At first sight, the problem is simple: If we know that TALYS + libraries are usually within e.g. 30% of the experimental data for a certain reaction channel, alarm bells should start ringing if the deviation of a data set for such a channel is suddenly much larger. However, large deviations may also come from bad performance of the library or TALYS, and may even occur if the visual agreement on linear scale is good. For example, for reactions close to threshold the difference between evaluation and experiment may easily be a factor of 10. In general the rule holds that the smaller the cross section, the larger the relative error. It is therefore important to judge not only the calculation/experiment (C/E) values and the reported experimental uncertainty, but also the absolute deviation. In several cases, it turns out that there are problems in EXFOR, and many of them can not so easily be detected with ways other than by comparing

with a model code, which is why these EXFOR problems have not been discovered in the first place. Therefore, this first global attempt to classify EXFO problems in a consistent manner is timely. The EXFOR mistakes which are easiest to detect concern C/E values around 0.001 or 1000, suggesting the well known error of mistaking barns for millibarns. Unfortunately, the majority of cases is more difficult to judge. The current comparison should also help to solve one category of mistakes in EXFOR: reaction identifiers which are assigned in wrong, inconsistent or even multiple ways, which can be regarded as an “injustice” for otherwise good-quality experimental data, since they are wrongly coded or sometimes even untraceable. In other words, if TALYS + libraries are expected to give a reasonably good prediction for a reaction and we obtain a large deviation, it may be that we are not comparing the evaluated result with the correct experimental quantity, and the EXFOR reaction identifier should perhaps be corrected.

To discover and classify problems, we first define 3 goodness-of-fit estimators, which represent an average deviation between TALYS or data libraries and an experimental data set that runs over N data points. If they are all 3 very large, something is wrong somewhere. They are the F-factor

$$F = 10\sqrt{\frac{1}{N}\sum_i^N f_i} \quad (1)$$

where the term for each individual data point is

$$f_i = \left(\log \frac{\sigma_T^i}{\sigma_E^i}\right)^2, \quad (2)$$

the χ^2 estimator,

$$\chi^2 = \frac{1}{N}\sum_i^N \chi_i^2, \quad (3)$$

where the term for each individual data point is

$$\chi_i^2 = \left(\frac{\sigma_T^i - \sigma_E^i}{\Delta\sigma_E^i}\right)^2, \quad (4)$$

and the absolute deviation

$$D = \frac{1}{N}\sum_i^N d_i, \quad (5)$$

where the term for each individual data point is

$$d_i = |\sigma_T^i - \sigma_E^i|. \quad (6)$$

In these equations, the subscript T stands for TALYS or libraries and E for experimental. Hence, each data set that contains a cross section excitation function, or only 1 point, is described by 3 average numbers: F , χ^2 and D , while in our analysis we also keep track of all individual components 10^{f_i} , χ_i^2 and D_i .

The F-factor is a kind of twisted C/E= σ_T/σ_E value. In fact, each f_i component of the sum inside F contributes to C/E if it is larger than 1, and E/C if it is smaller than 1. This is a more

appropriate quantity than the average C/E, since averaging C/E values over many points may not be very meaningful if the individual values cross unity at some point. Eq. (1) remedies this. A value of $F=1.2$ means that for the entire data set we are roughly 20% off on average. We use F as the leading indicator in our statistical study, i.e. we sort our results in order of increasing F to identify the worst cases. Note that the term “data set”, i.e. the sum over N , can apply to one EXFOR subentry, e.g. one excitation function, all subentries for the same (Z,A) nuclide and reaction channel (MT), all subentries for the same (Z,A) nuclide, all subentries for the same reaction channel (MT), all subentries for the same projectile, and finally to the entire part of the EXFOR database that could be compared. For all this, we have average F -values. In addition, all these averages can be taken for each library or TALYS separately, or summed over all of them. For the purpose of checking EXFOR, the goodness-of-fitness for one subentry, i.e. one experimental data set for 1 or a range of incident energies, is used.

Another standard indicator is of course χ^2 , and then the extra aspect for quality testing is that apart from the central values the uncertainties given in EXFOR need to be reliable as well. If χ^2 does not have the same rank as F , there may be something wrong with the reported uncertainty.

Finally, large F or χ^2 values may actually be acceptable, if the underlying quantities have a small value. This occurs often for measurements close to threshold. To identify those cases, the absolute deviation in mb, D is an additional helpful quantity.

In sum, for quality testing of EXFOR it is best to monitor all three indicators simultaneously.

Now that we have defined the goodness-of-fit estimators, we can give the current quality classification:

- T1, N1, R1 and E1: $1 \leq F \leq 1.2$
- T2, N2, R2 and E2: $1.2 < F \leq 2$.
- T3, N3, R3 and E3: $F > 2$.

At the moment of this writing, we feel that quality class '3' should always be reviewed by checking the associated documentation (publication), while class '2' should be reviewed if D is more than 10% of the total nonelastic cross section and if χ^2 is larger than 20.

Clearly, when testing EXFOR we need to be clever. Efficiency is maximal if we can minimize:

- False negative: the goodness-of-fit estimators suggest there is no problem, while in reality there is. The problem remains in EXFOR.
- False positive: the goodness-of-fit estimators suggest there is a problem, but checking the publication reveals that the correct data is entered into EXFOR. We may have wasted our time.

As in real life, we may be more worried about false negatives than false positives, so it may be wise, time permitting, to take a conservative approach towards a quality classification scheme. Actually, a false positive at least gives the possibility to assign a 'suspicious quality' label to the corresponding data set, since the compiler made no mistake, but the goodness-of-fit estimators are well away from the average. In an attempt to use all three goodness-of-fit estimators at once, we are currently investigating a new set of estimators, the K-factors, which should be more full-proof to identify whether a data point or set is "wrong". A K-factor for a data set of N data points is defined as

$$K^{(k)} = 10\sqrt{\frac{1}{N}\sum_i^N k_i^{(k)}}, \quad (7)$$

where $k_i^{(k)}$ is the goodness-fit-estimator per point. The "0-th order" K-factor is just the F-factor, i.e.

$$k_i^{(0)} = f_i = \left(\log \frac{\sigma_T^i}{\sigma_E^i}\right)^2 \longrightarrow K^{(0)} = F \quad (8)$$

To minimize the false alarms for a wrong data set, we introduce various "penalties" into the definition of K , so that the K-distribution becomes more stretched than the F-distribution, and "true" outliers deviate even further from unity.

First, the weight of the absolute distance D can be taken into account by comparing the theoretical, or evaluated, result with the theoretical nonelastic cross section σ_{non}^i . Threshold problems or (very) weak channels are then inhibited, which for quality testing is justified since it is rather common to have a large relative deviation between theory and experiment, without one the two being clearly wrong. For each data point we define a factor

$$r_i = \left(1 + \frac{\sigma_T^i}{\sigma_{non}^i}\right)^{p_r} \quad (9)$$

where p_r is a power to scale the importance of this case. Currently, $p_r = 2$. The goodness-of-fit estimator $K^{(1)}$ is then given by Eq. (7) with

$$k_i^{(1)} = f_i \cdot r_i \quad (10)$$

Next, the weight of the experimental error is taken into account by comparing it with the deviation of the evaluation from the central value, i.e.

$$e_i = \left(1 + \frac{f_i - 1}{d\sigma_E^i/\sigma_E^i}\right)^{p_e} \quad (11)$$

where p_e is a power to scale the importance of this case. Currently, $p_e = 2$. With this factor, we hope to catch cases where possibly the reported precision of the measurement was too optimistic, or a wrong compilation of the experimental error. In any case, it will amplify the deviation from the norm. Taking this into account, the goodness-of-fit estimator $K^{(2)}$ is then given by Eq. (7) with

$$k_i^{(2)} = f_i \cdot r_i \cdot e_i \quad (12)$$

Actually, a third penalty can be introduced, namely the deviation of an experimental data point from the other experimental points for the same reaction. So far, all deviations can be seen as “theoretical outlier” or “evaluated outlier”, while comparing with similar measurements would lead to purely “experimental outliers”. For example, it could happen that there are many experimental data sets for the same reaction with an F between 1 and 1.10, while one measurement has $F=1.50$. This value itself would not give rise to alarm, but the fact that it deviates from so many other sets does. For this extra test a second loop over the whole database is needed, i.e. after it has been created, and this is not yet implemented.

While we make the translation from the XC4 computational database to our own directory-structured database, we do our TALYS + library comparison, checks and statistical analyses on the fly. After a few hours, the conversion is done and all checking and statistical results are available.

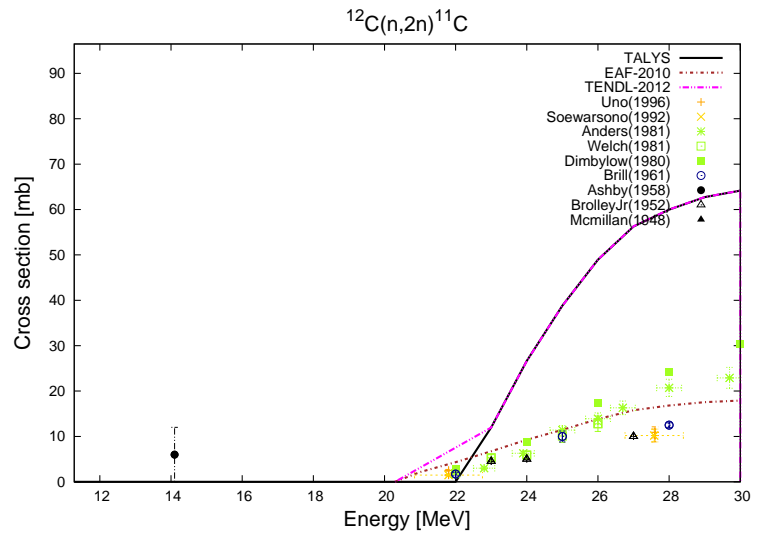
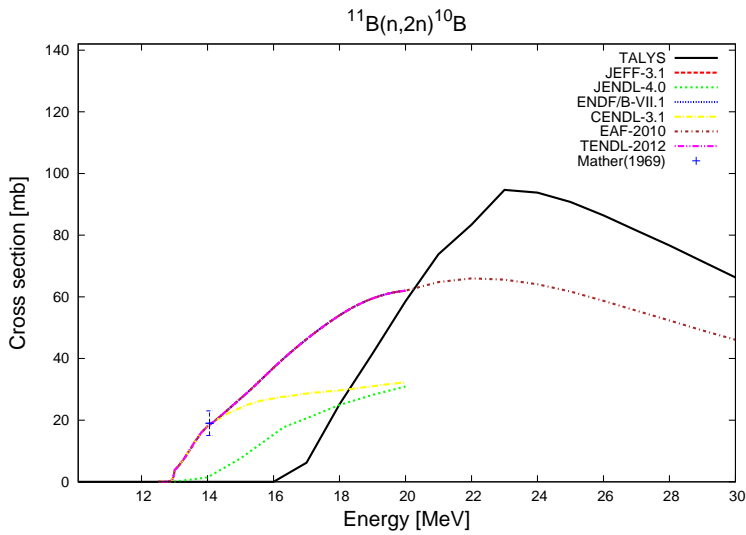
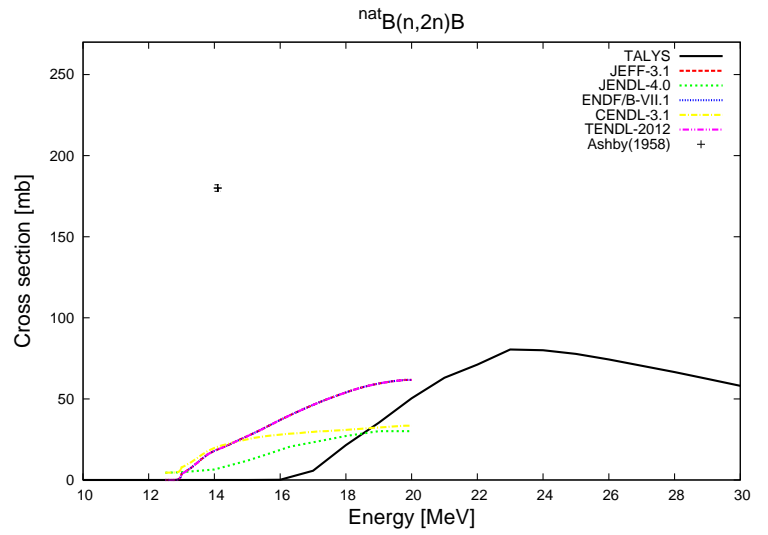
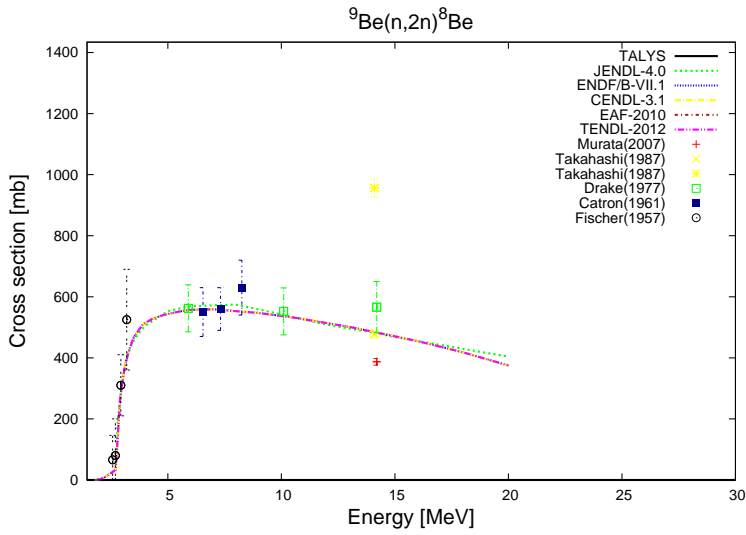
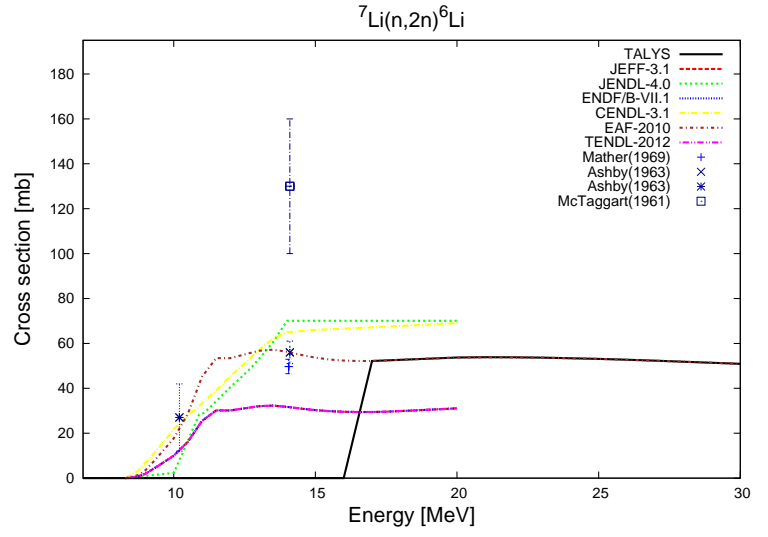
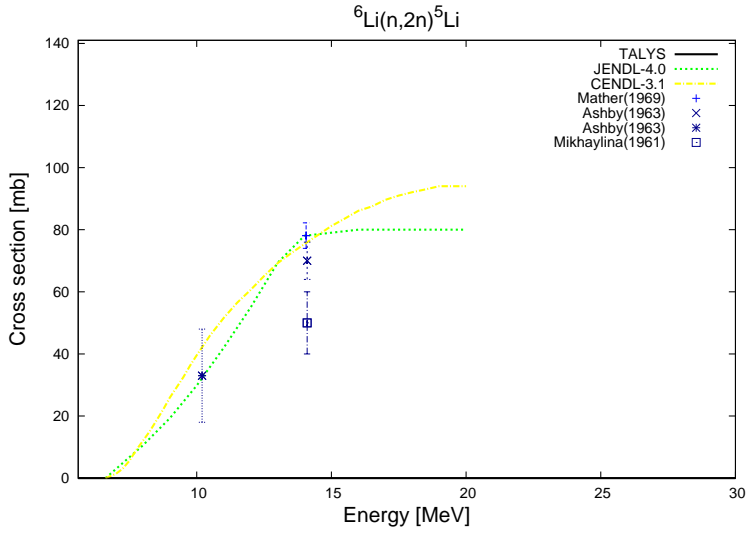
Score	(n,2n)	(n,p)	(n, α)
T1	684	320	153
T2	402	452	270
T3	100	95	127
R1	309	265	84
R2	292	454	179
R3	87	162	143
N1	495	277	116
N2	380	386	190
N3	80	120	115
E1	0	0	0
E2	4	5	1
E3	6	11	2

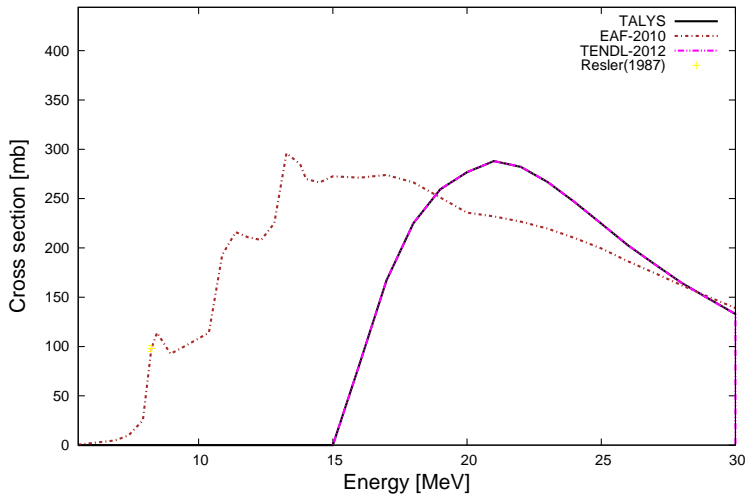
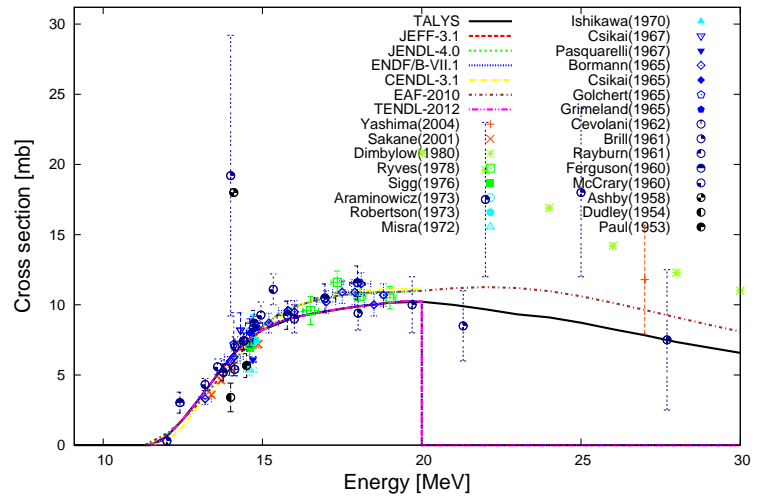
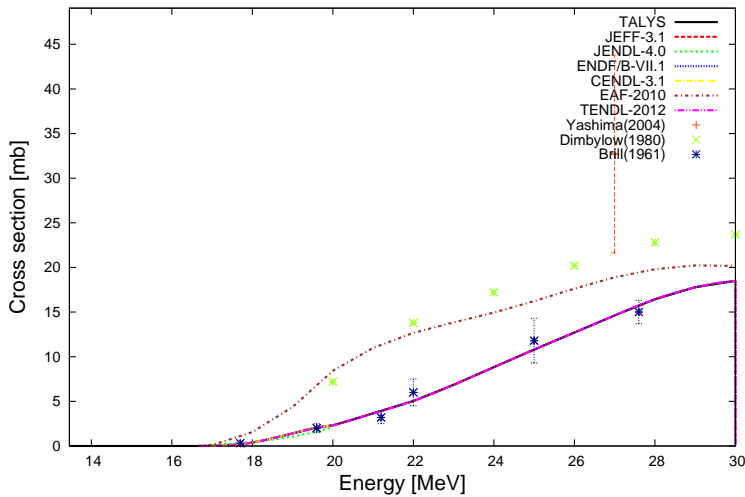
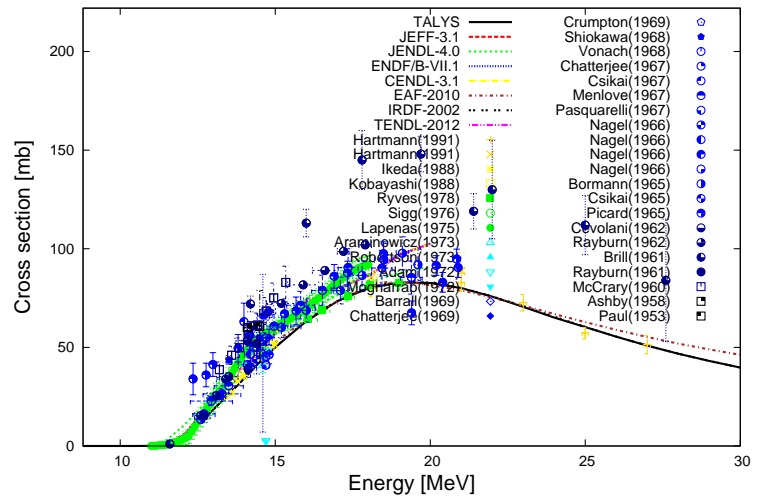
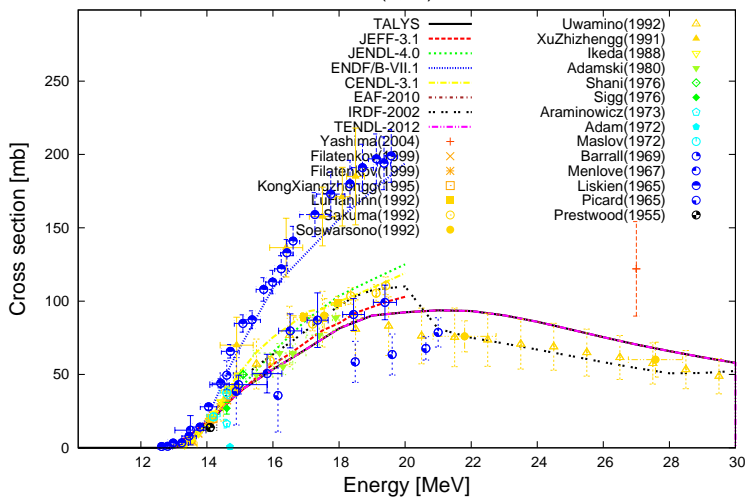
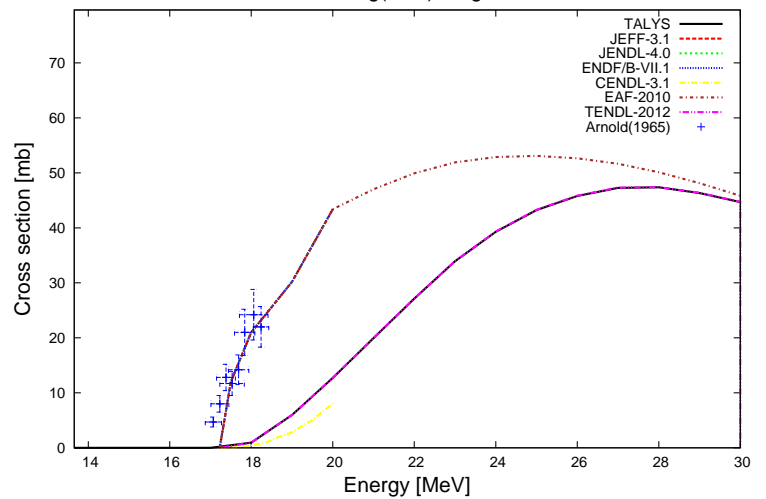
Table 2: Quality scoring of (n,2n), (n,p) and (n, α) scoring as of April 22 2013.

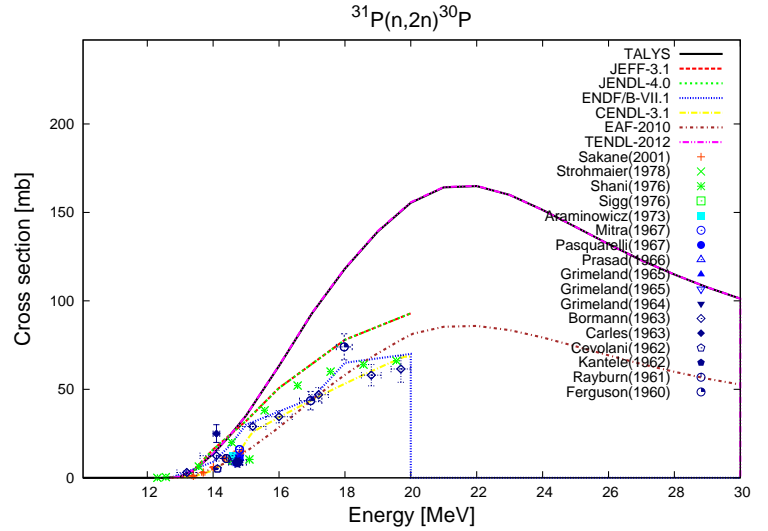
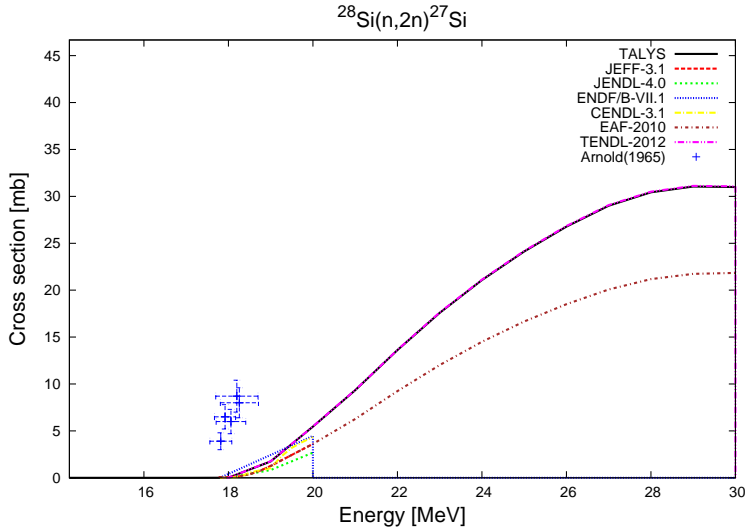
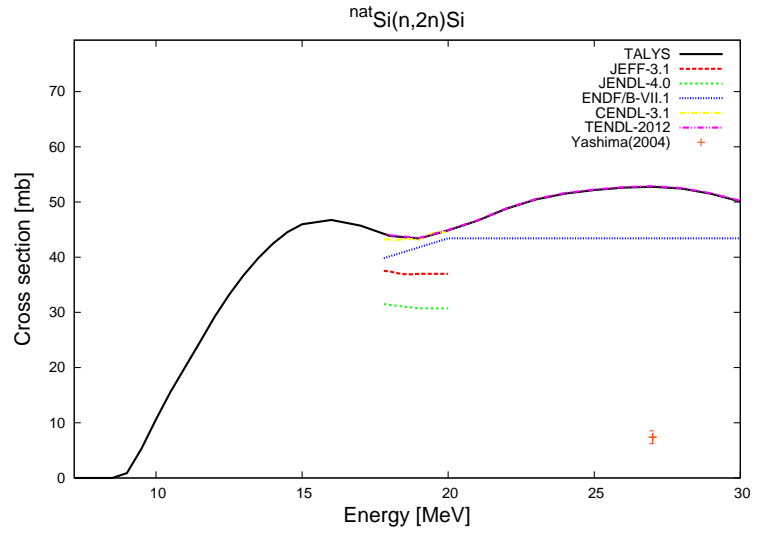
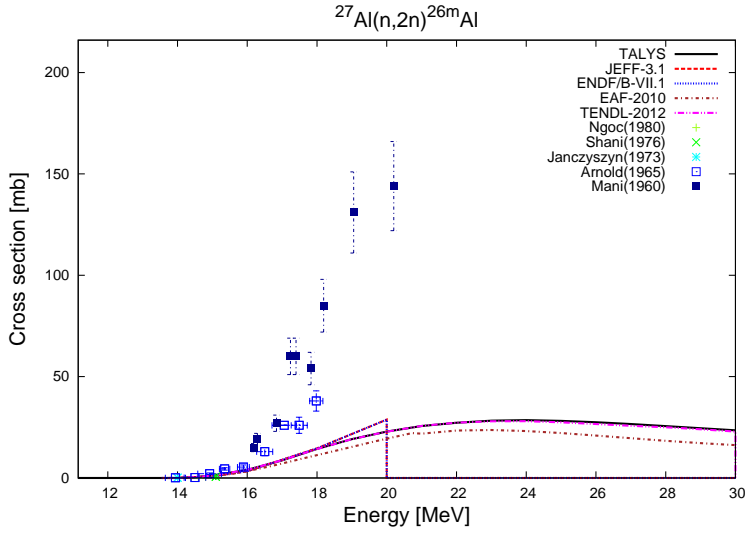
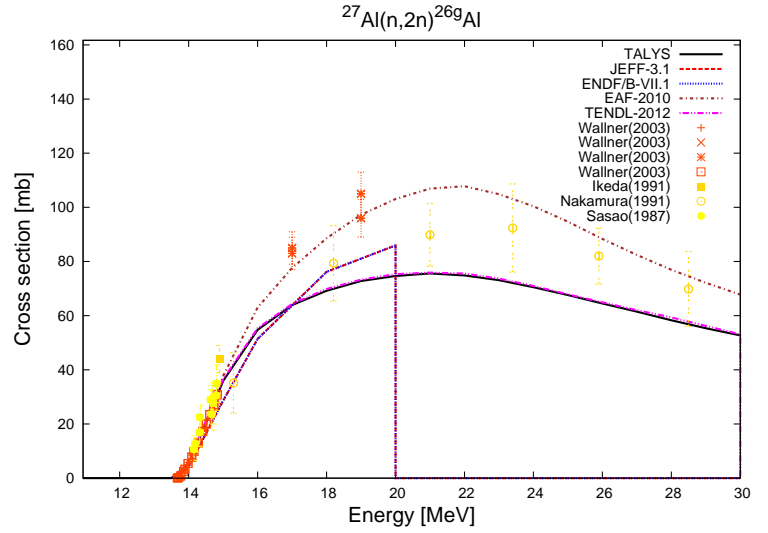
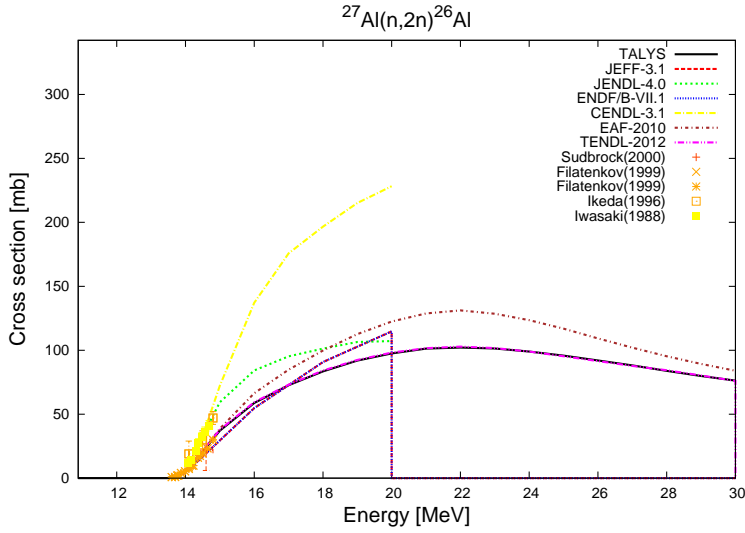
4 Graphical comparison

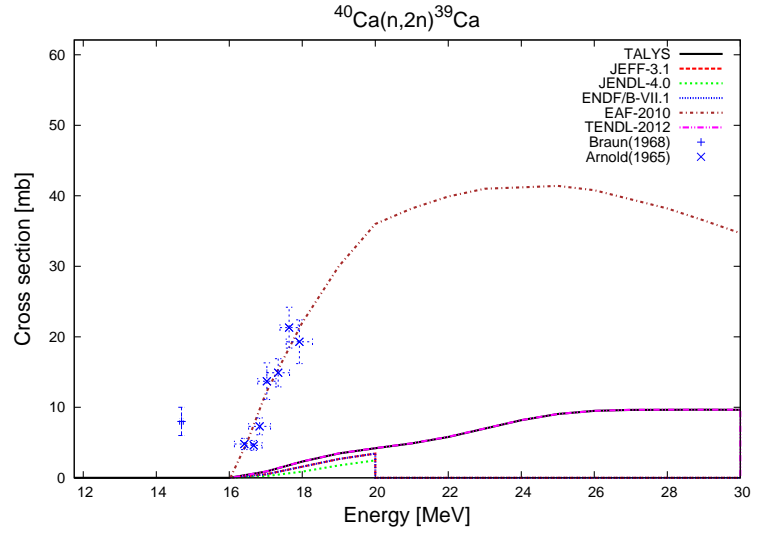
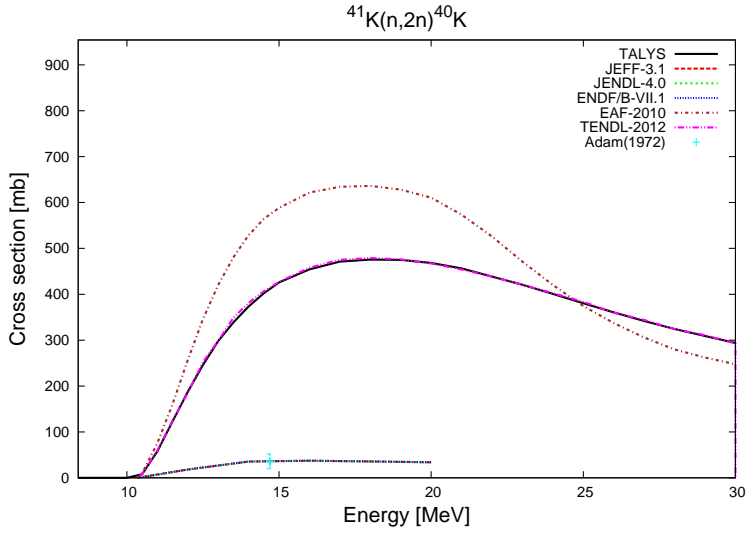
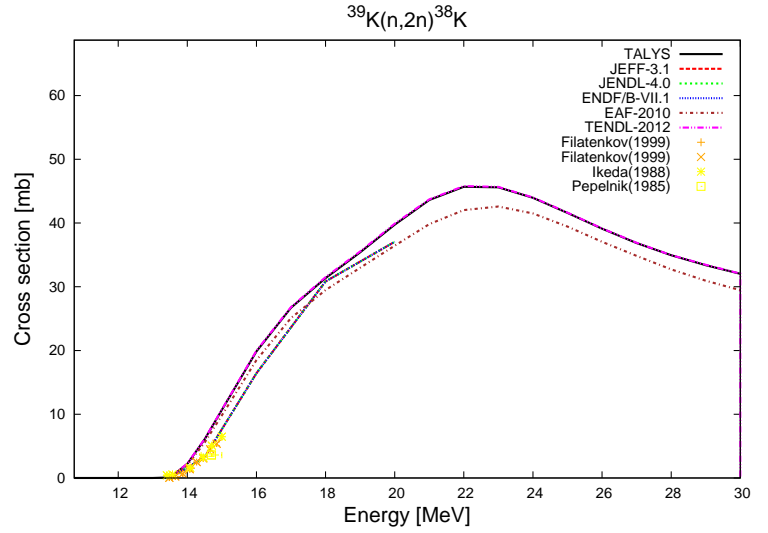
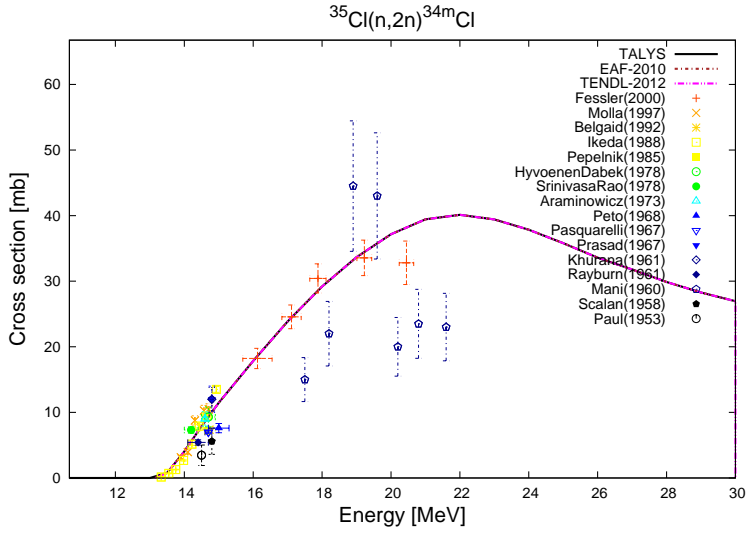
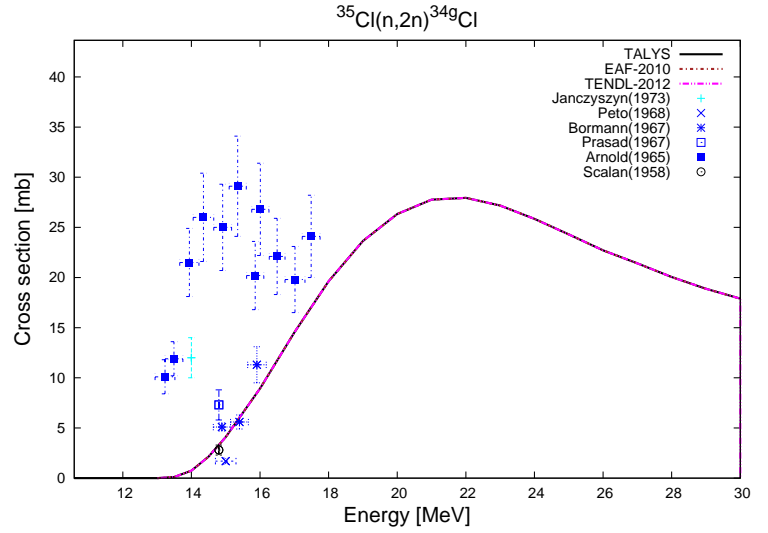
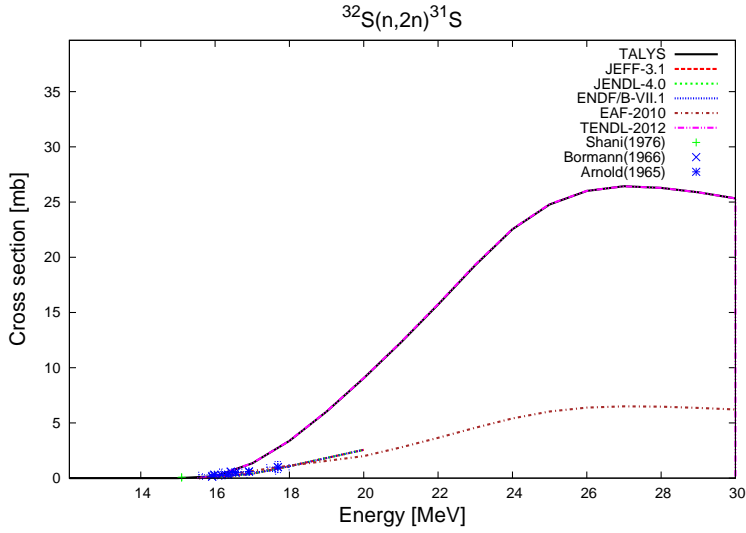
For each combination projectile - target nucleus - reaction channel, we produce an automatic plot in which all experimental data sets are compared with the most important nuclear data libraries and/or the latest TALYS calculations. This allows to test whether numerical outliers are also graphical outliers.

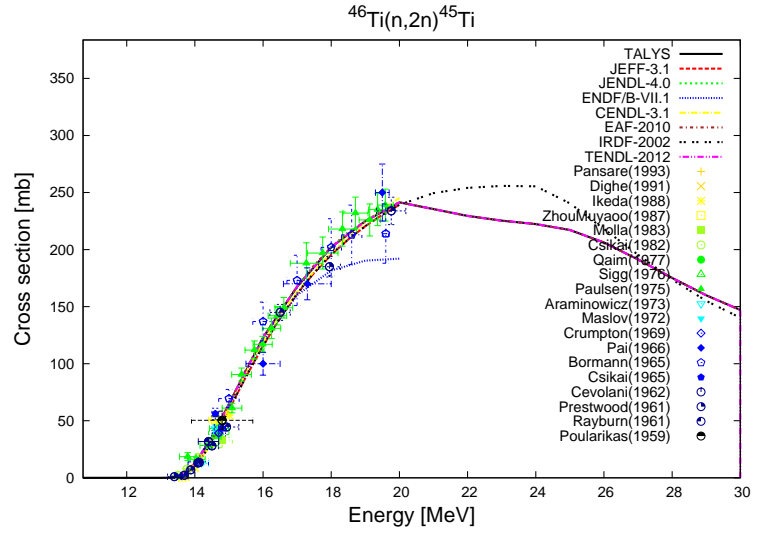
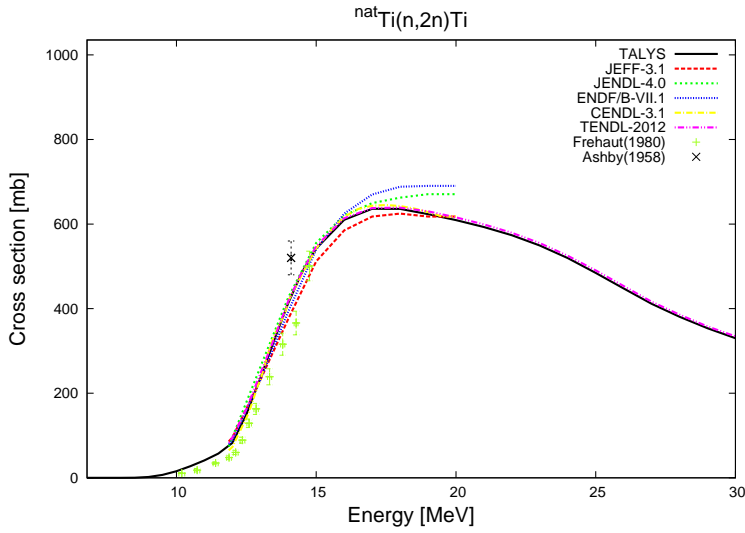
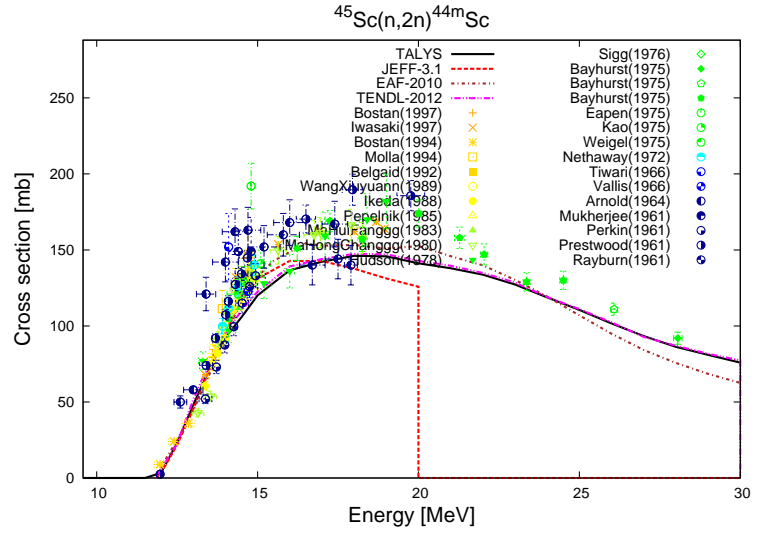
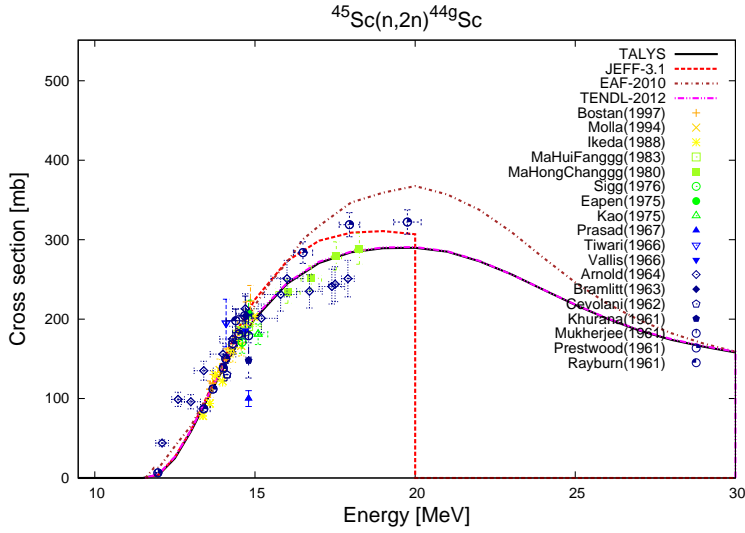
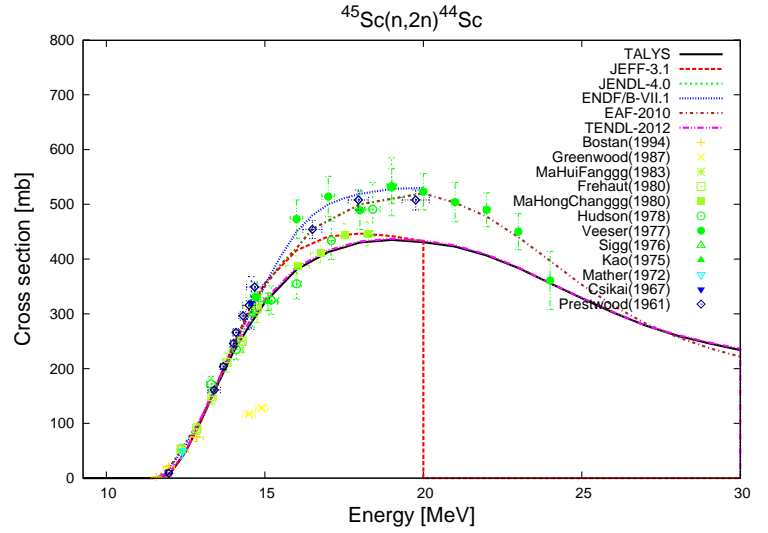
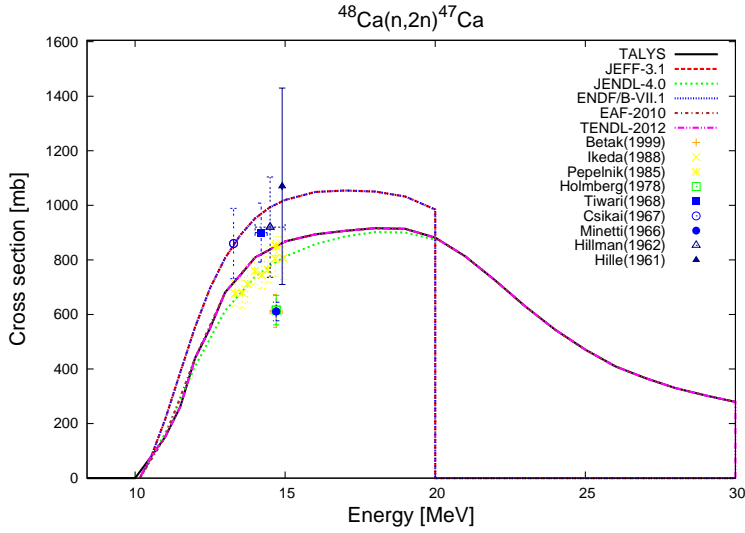
4.1 Activation cross sections: (n, 2n), (n, p) and (n, α)

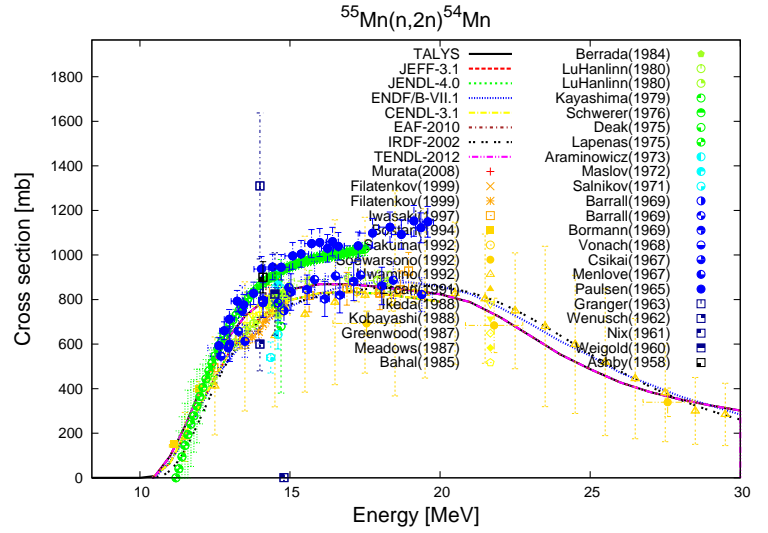
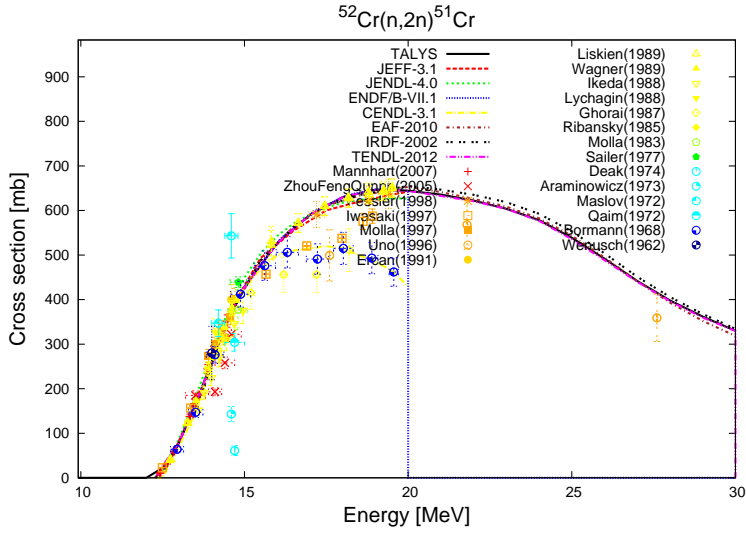
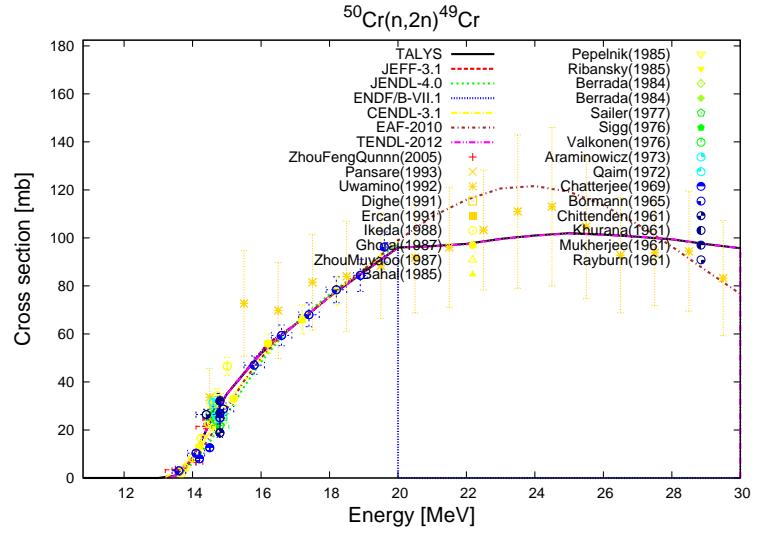
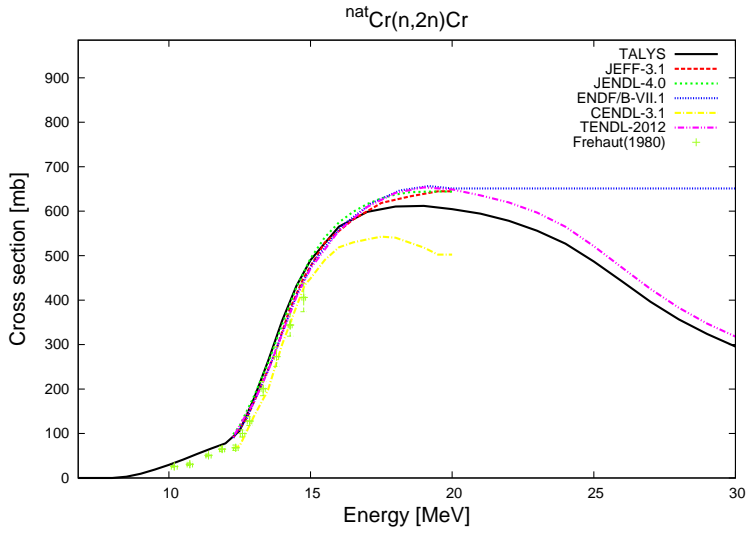
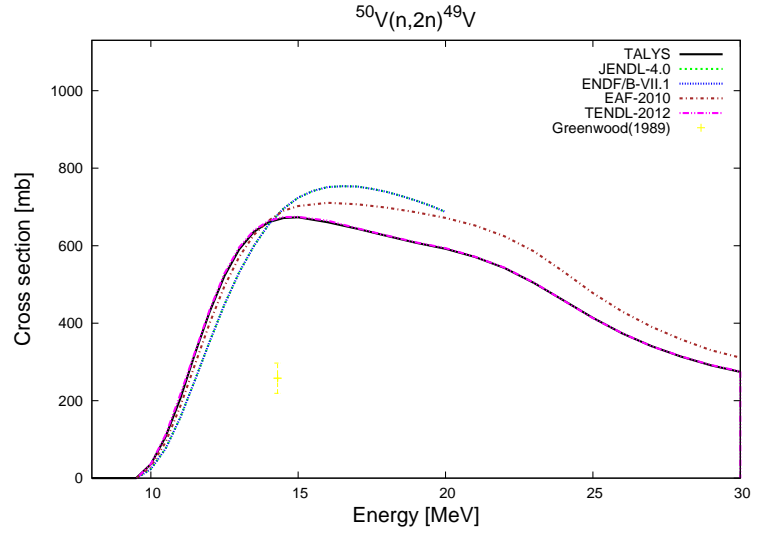
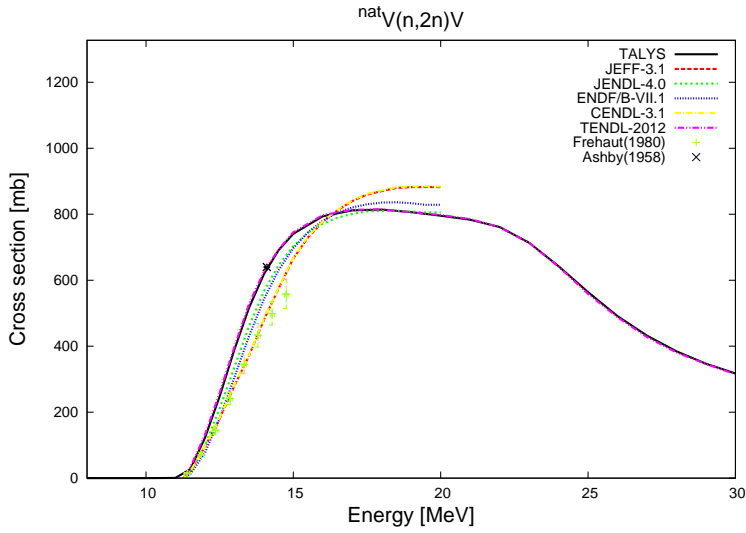


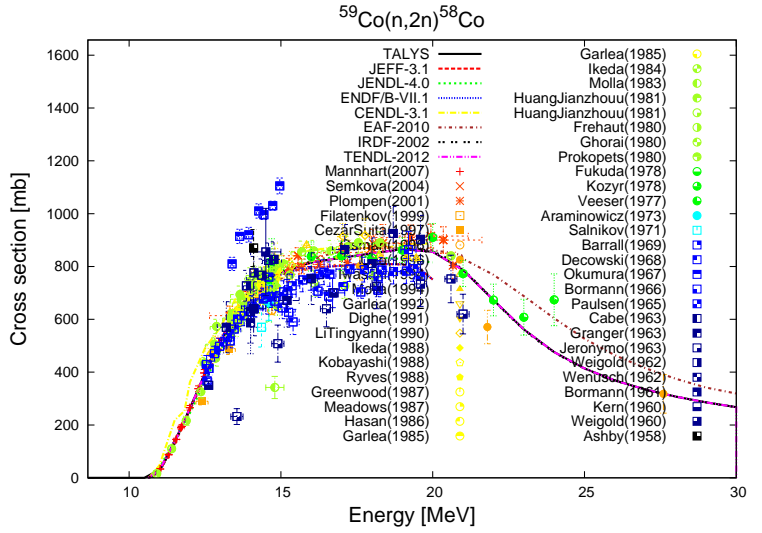
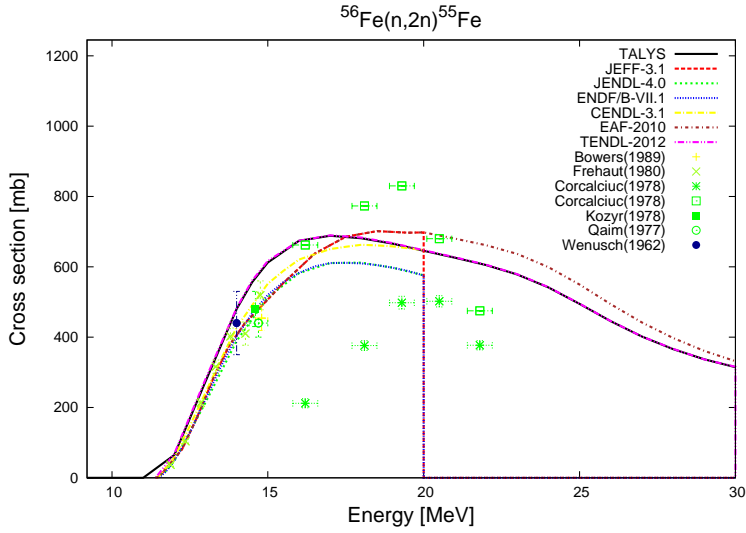
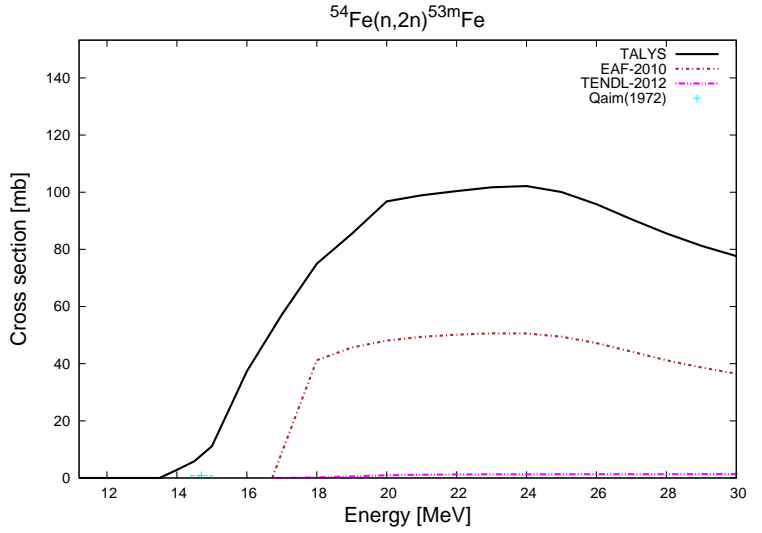
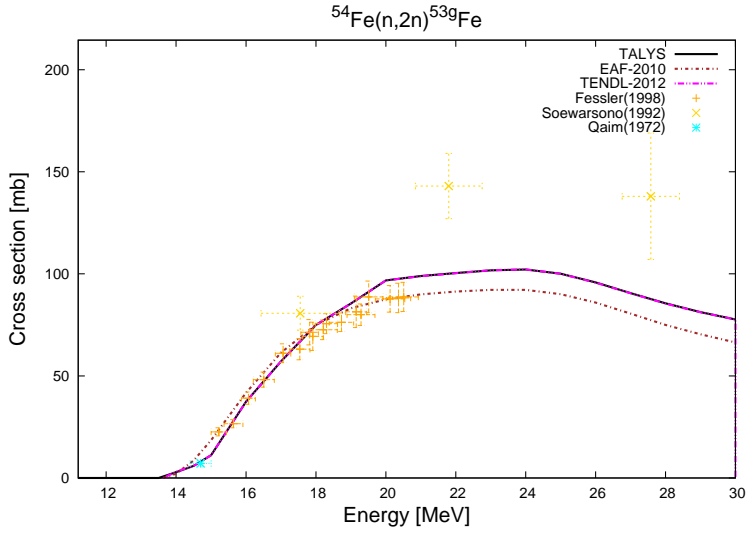
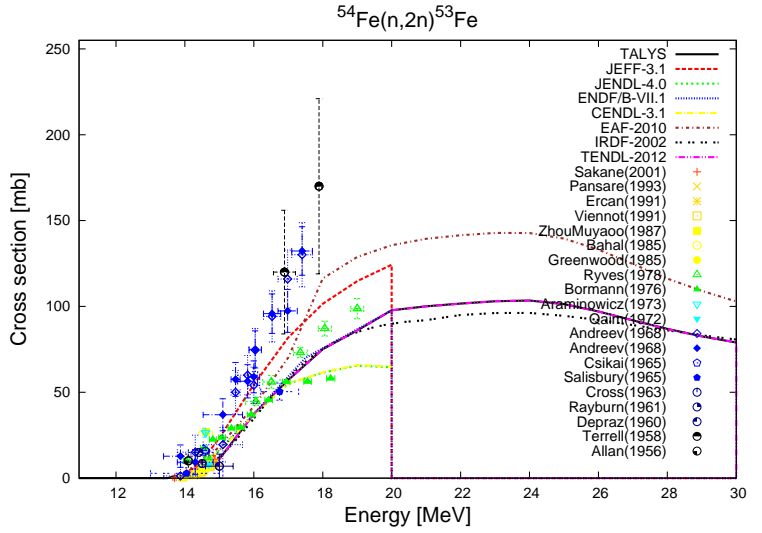
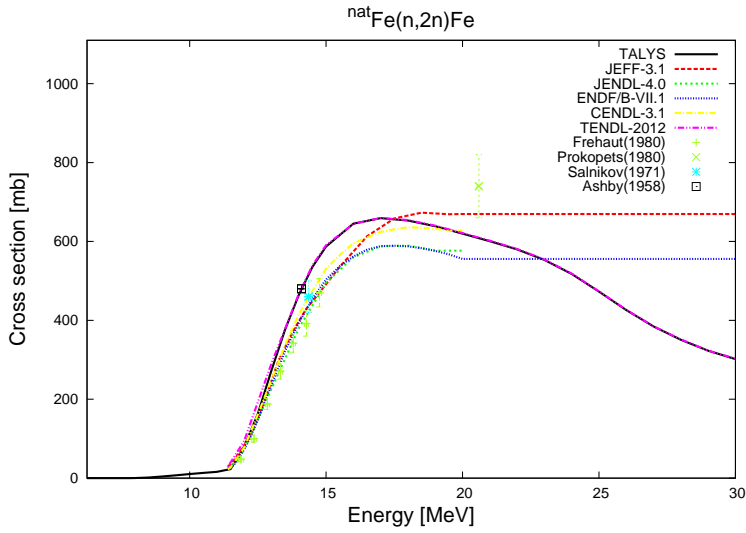
$^{13}\text{C}(n,2n)^{12}\text{C}$  $^{14}\text{N}(n,2n)^{13}\text{N}$  $^{16}\text{O}(n,2n)^{15}\text{O}$  $^{19}\text{F}(n,2n)^{18}\text{F}$  $^{23}\text{Na}(n,2n)^{22}\text{Na}$  $^{24}\text{Mg}(n,2n)^{23}\text{Mg}$ 

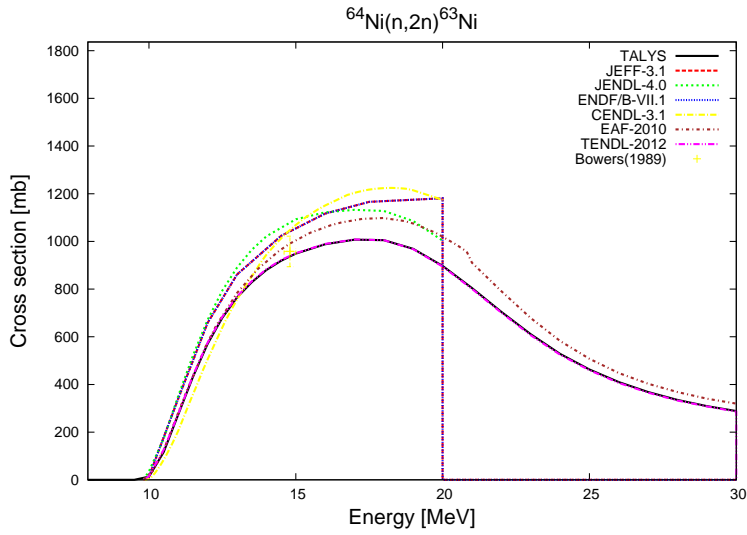
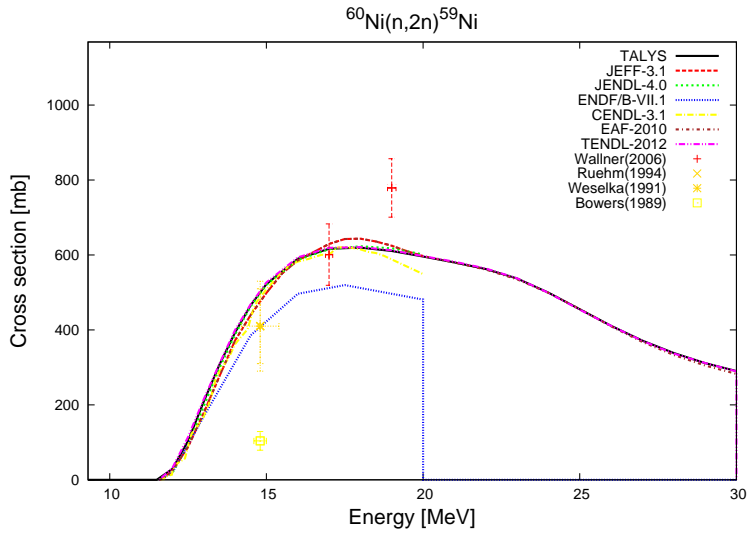
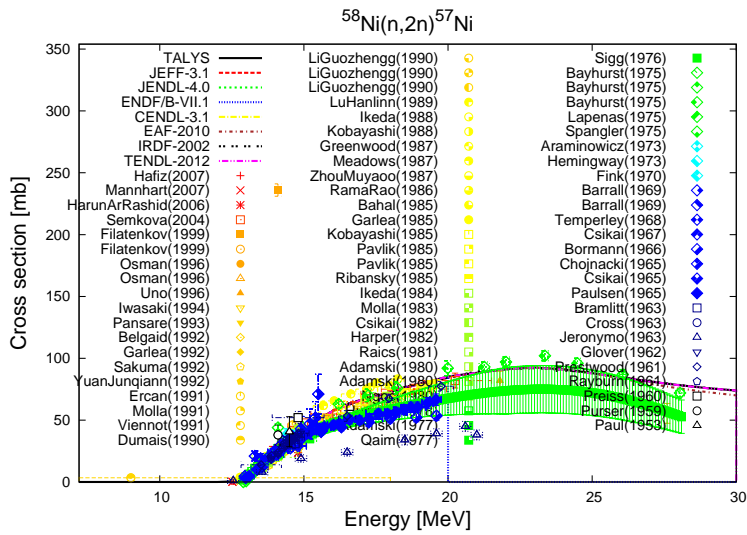
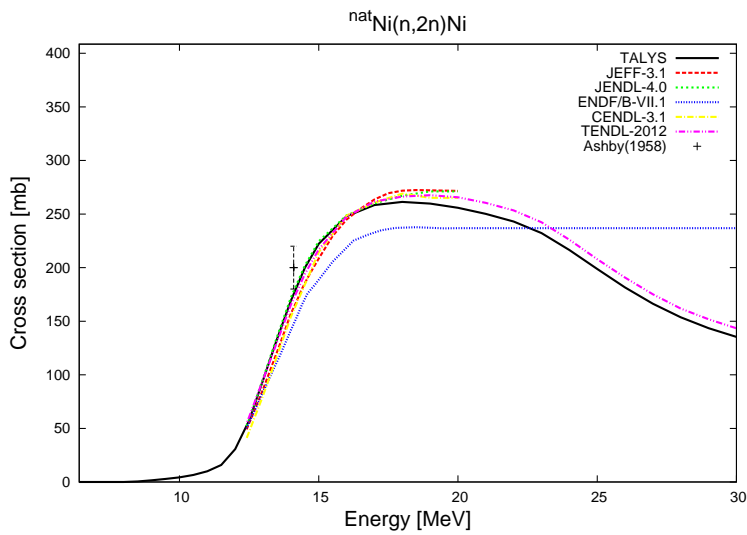
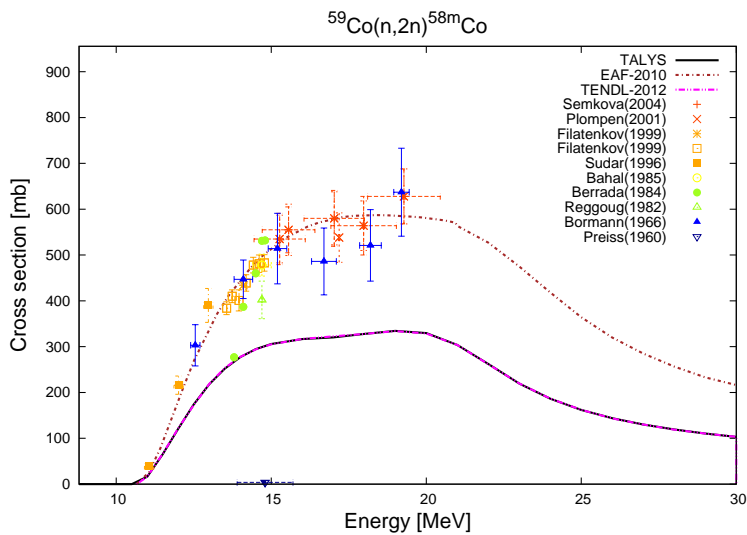
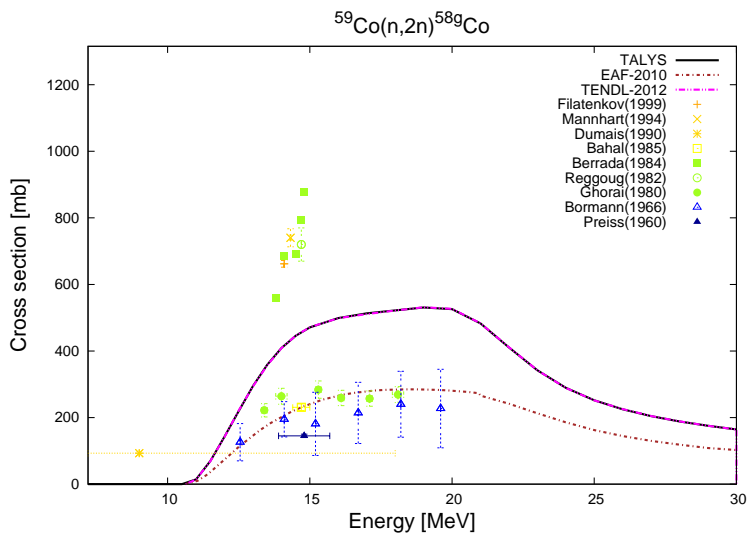


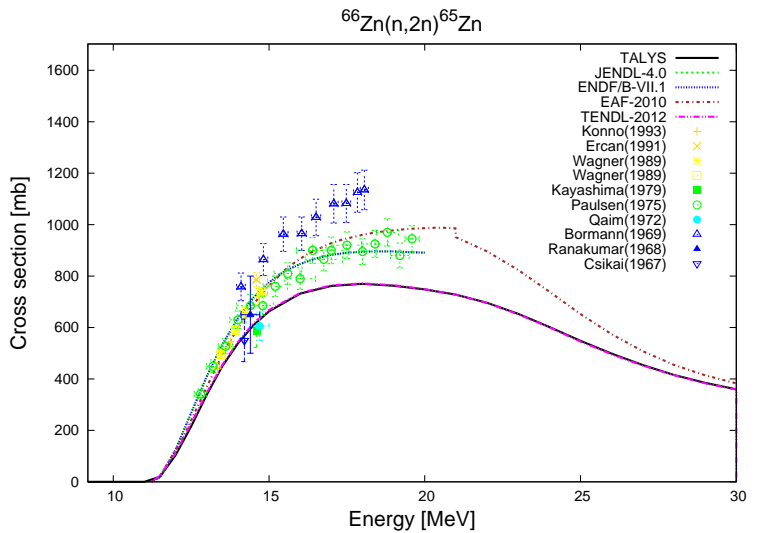
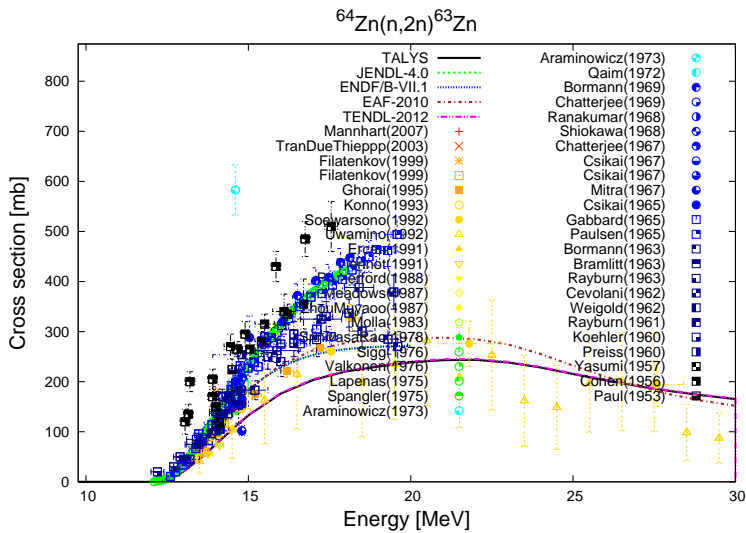
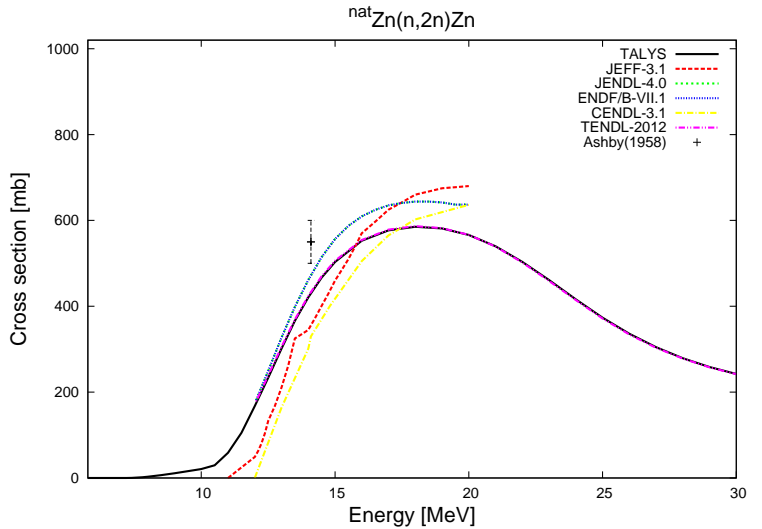
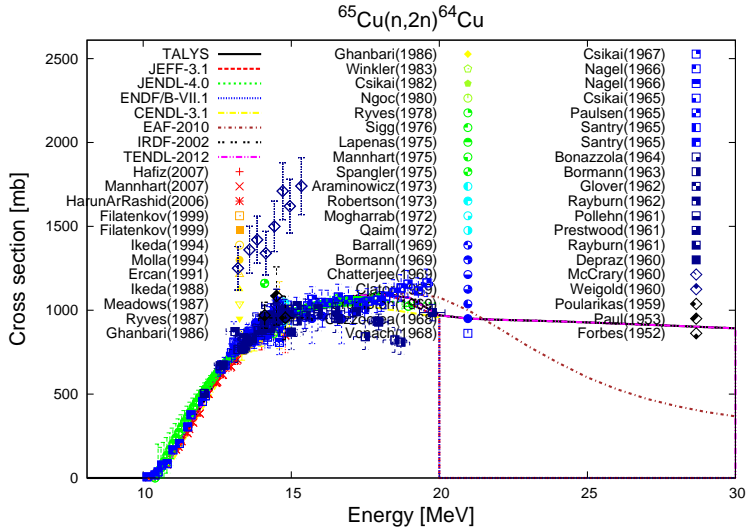
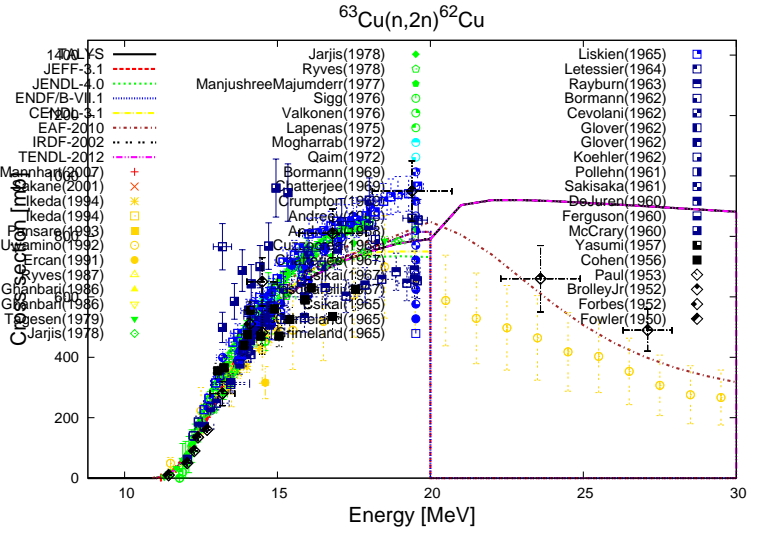
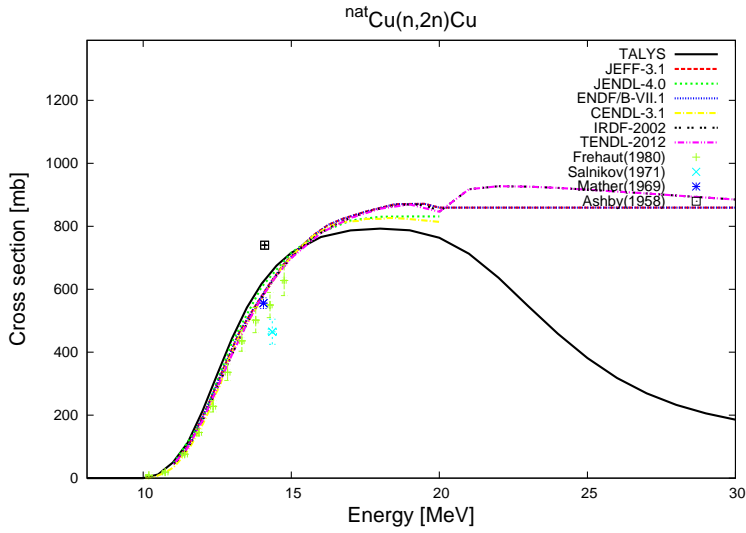


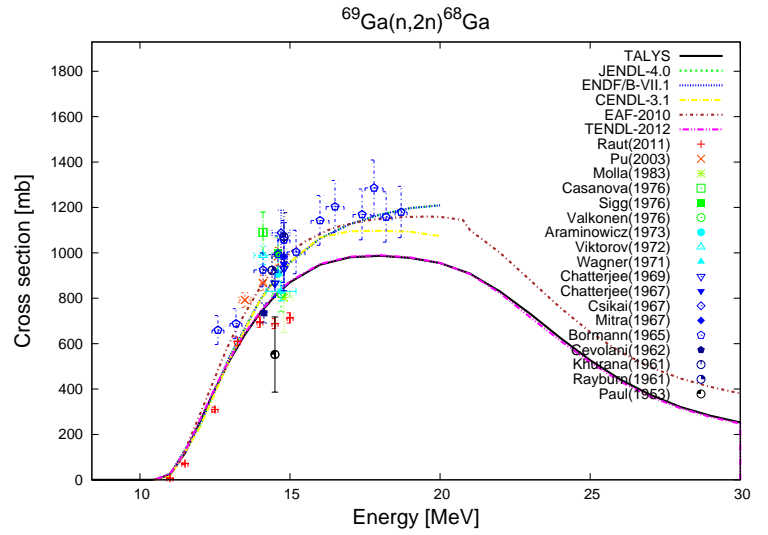
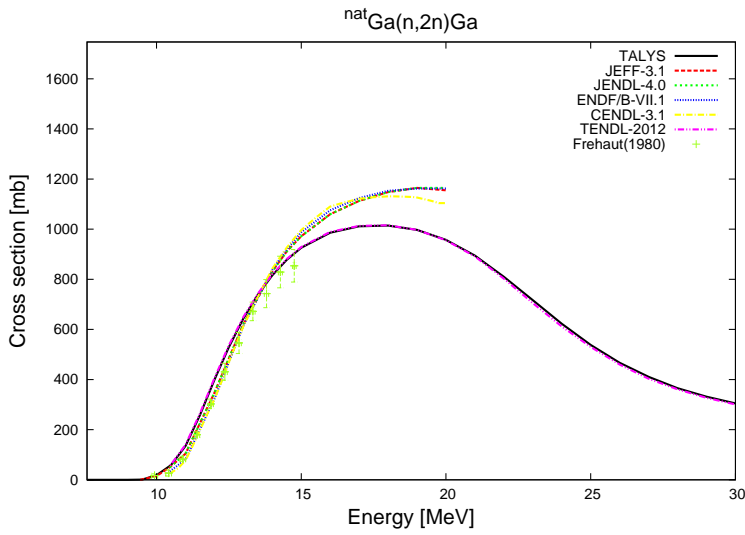
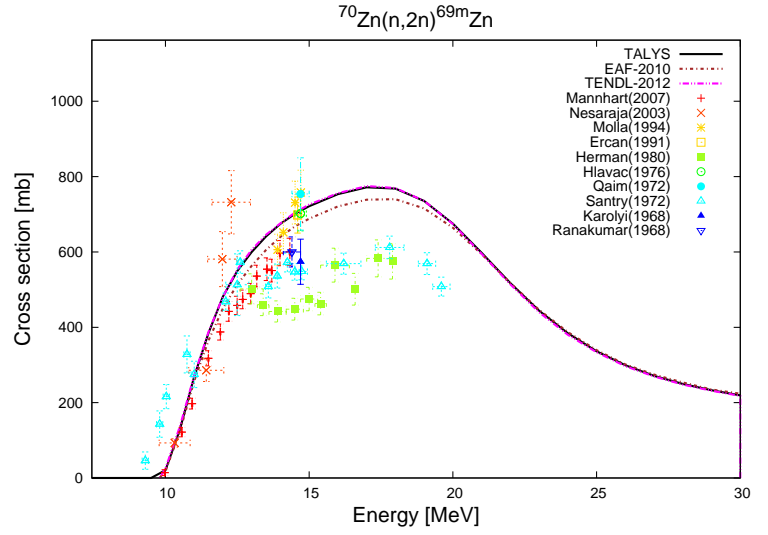
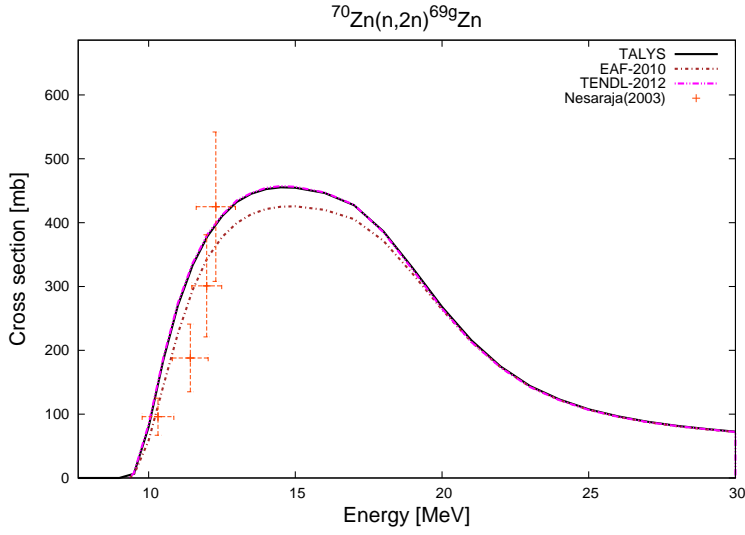
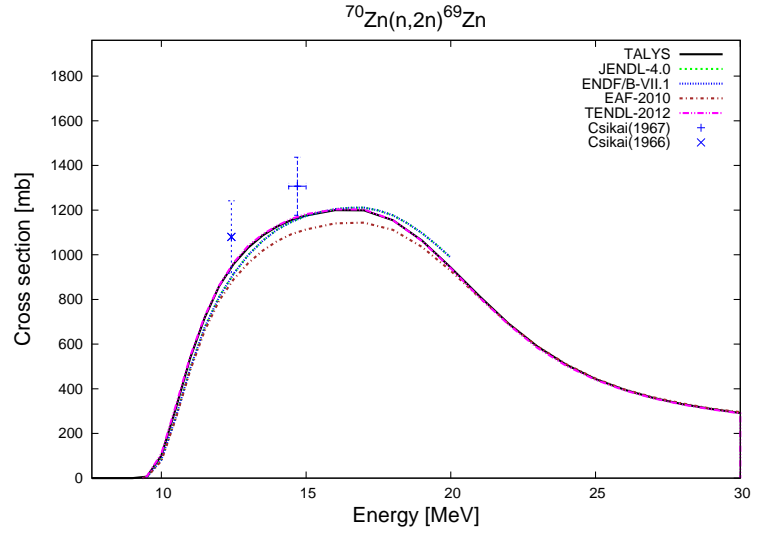
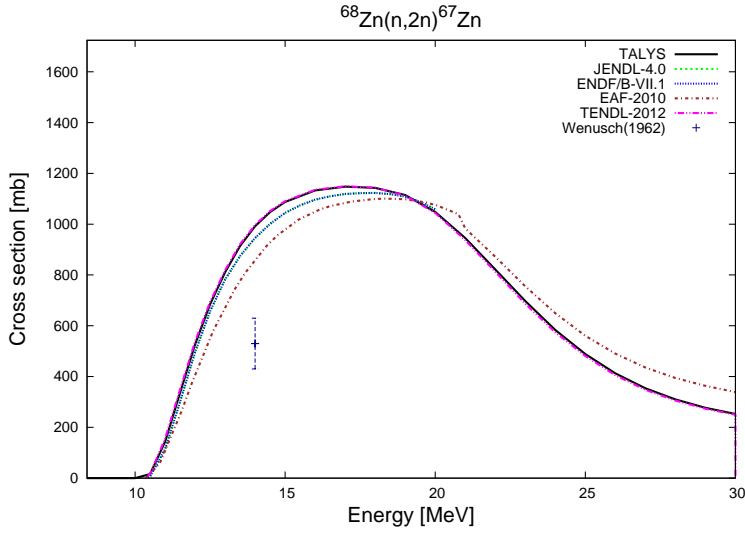


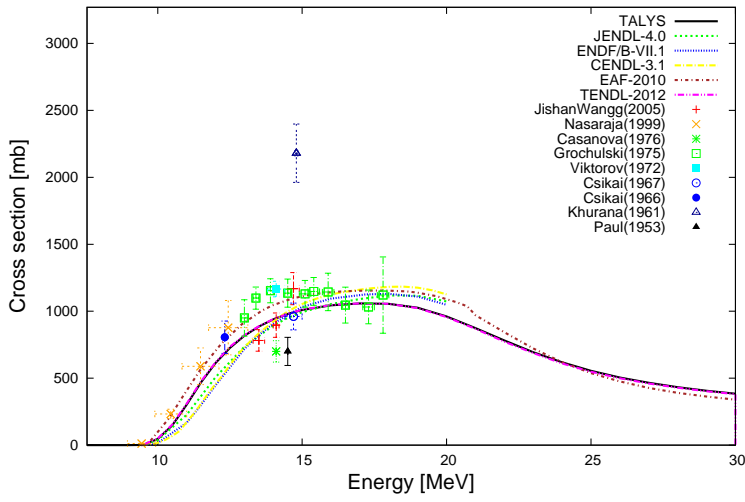
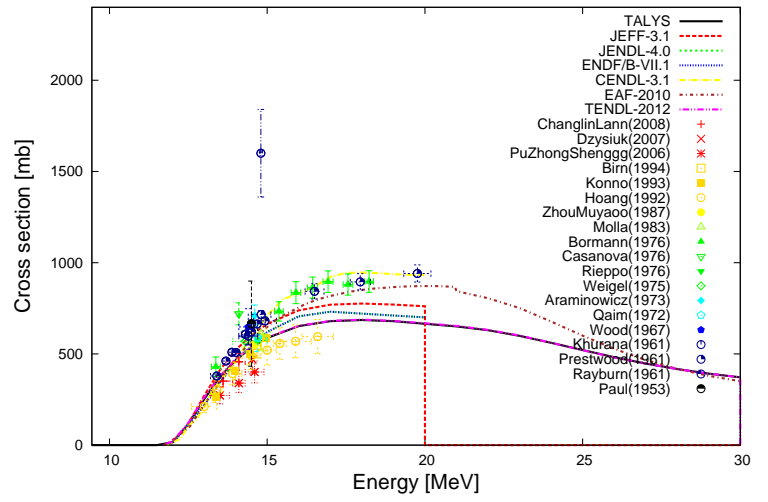
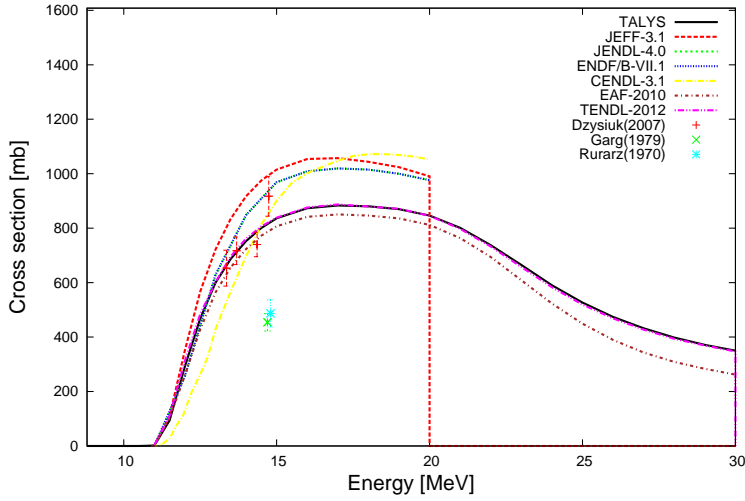
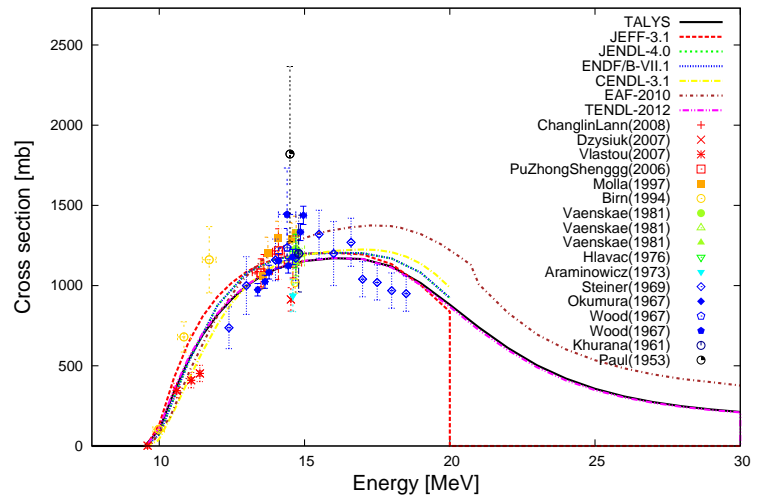
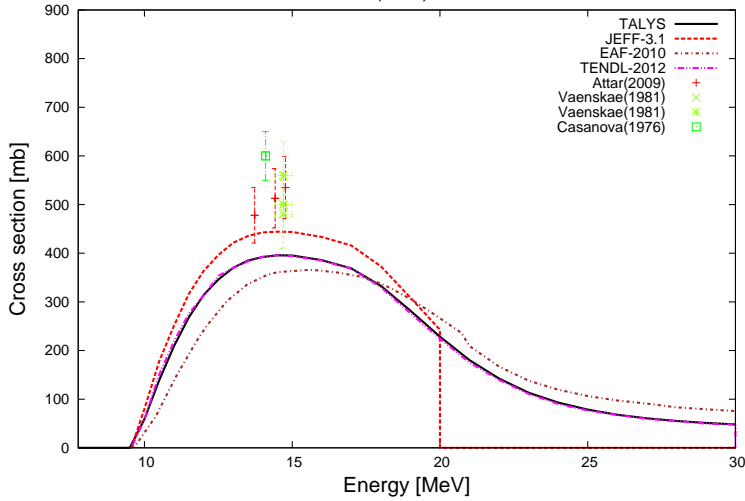
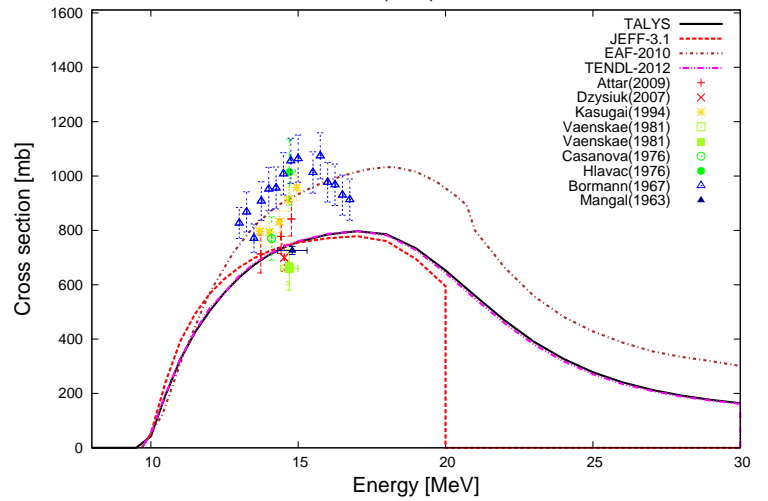


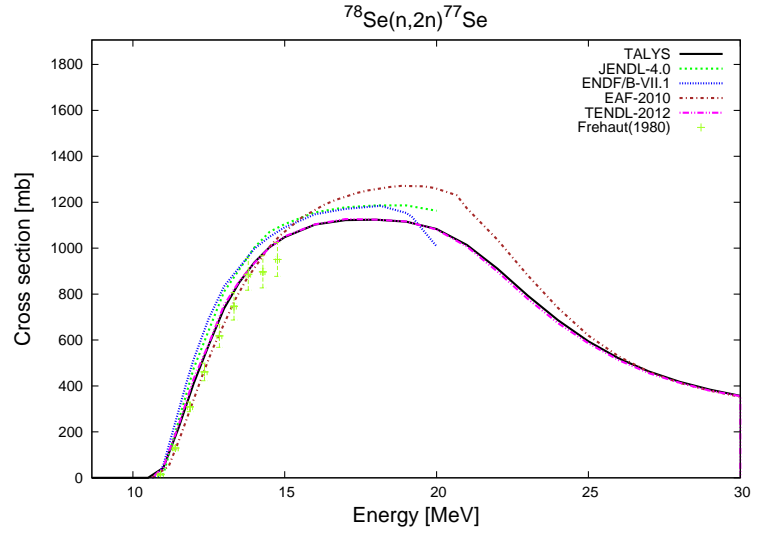
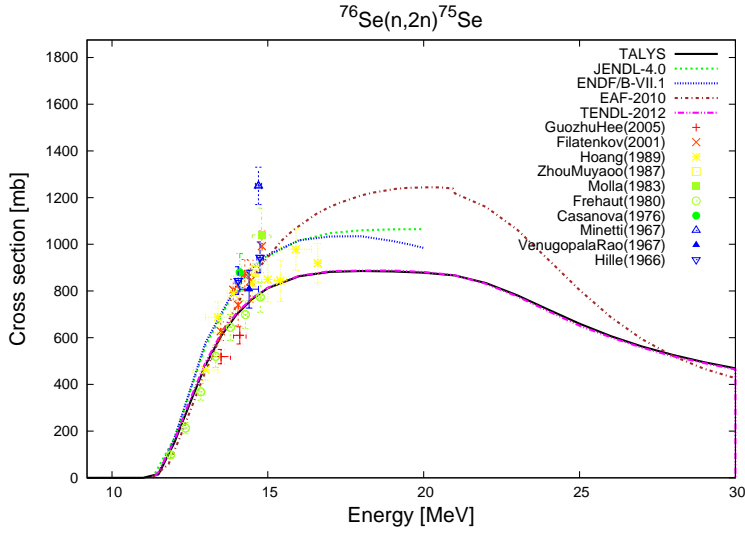
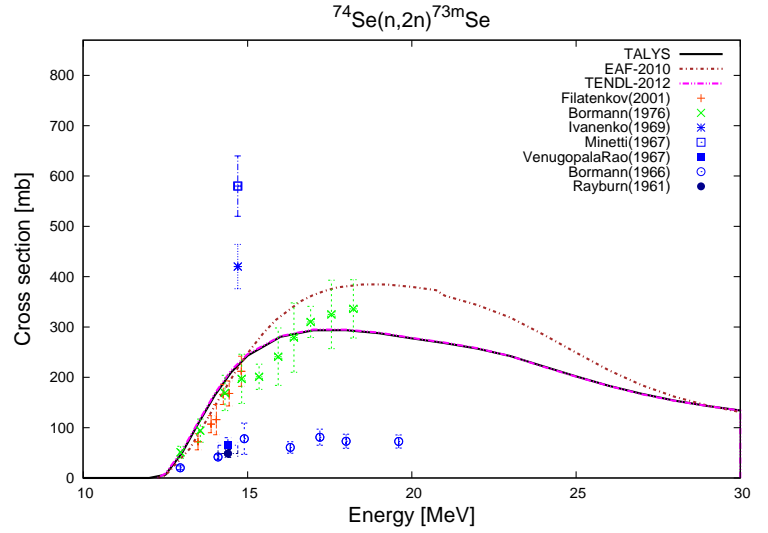
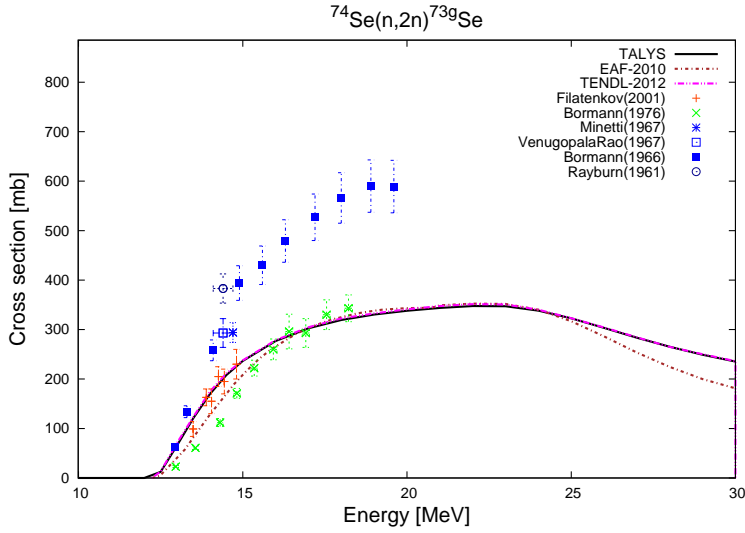
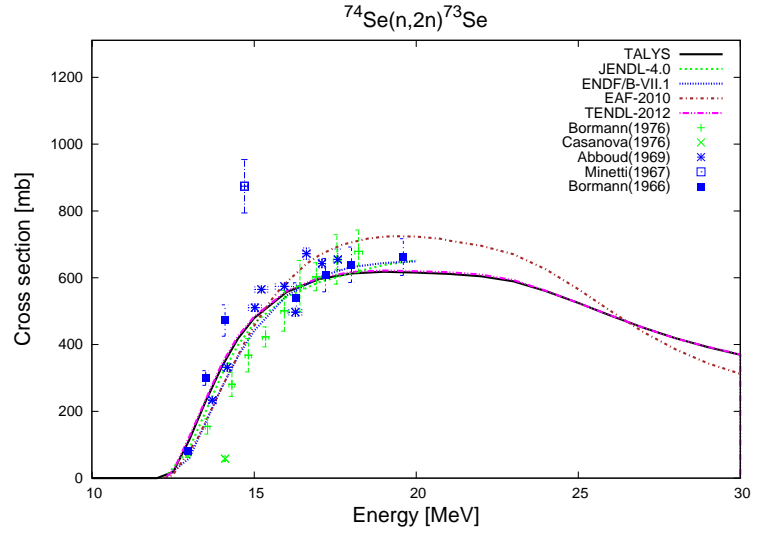
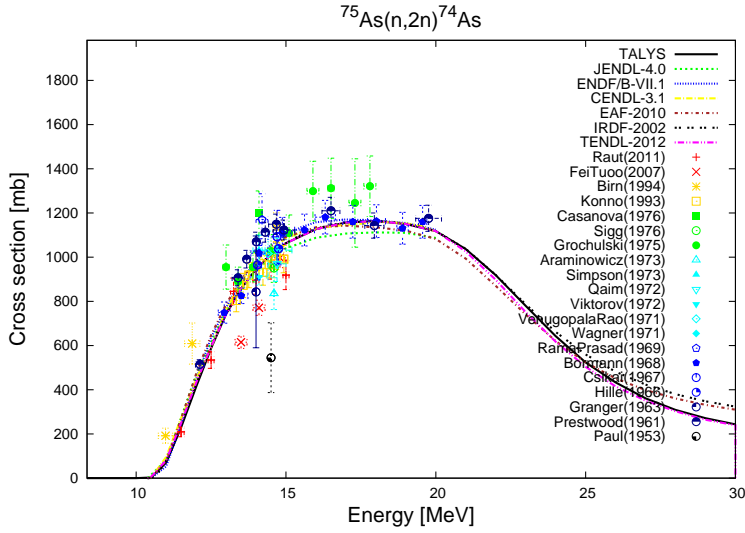


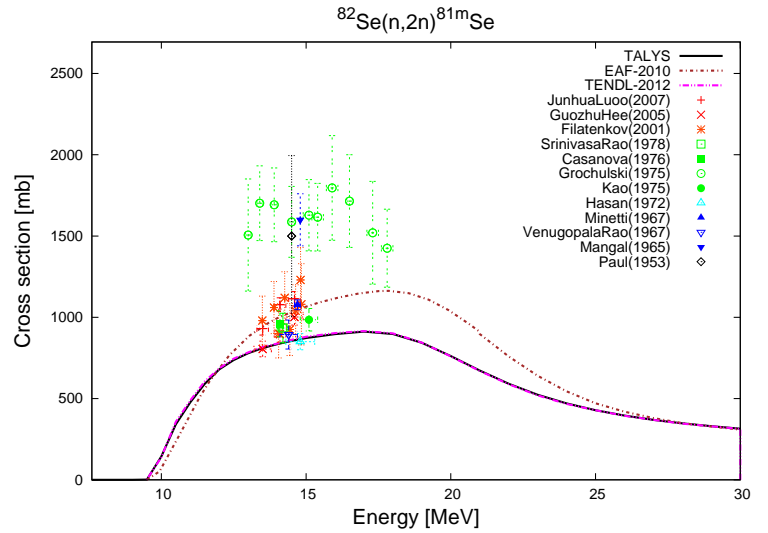
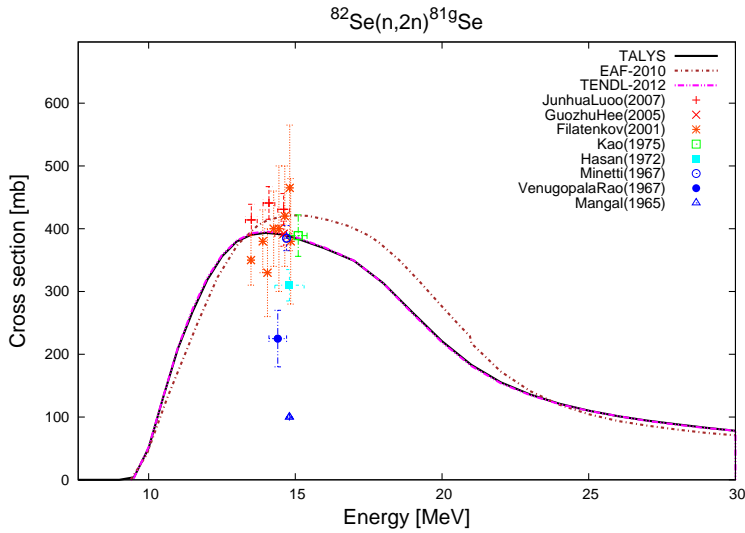
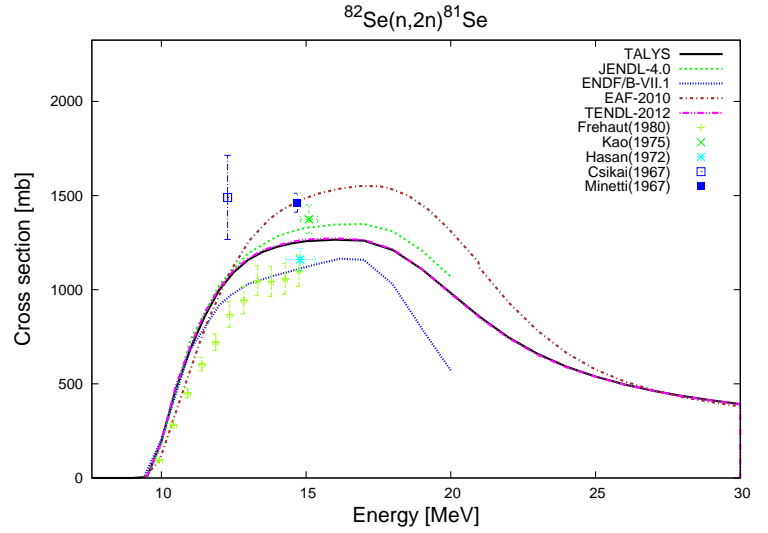
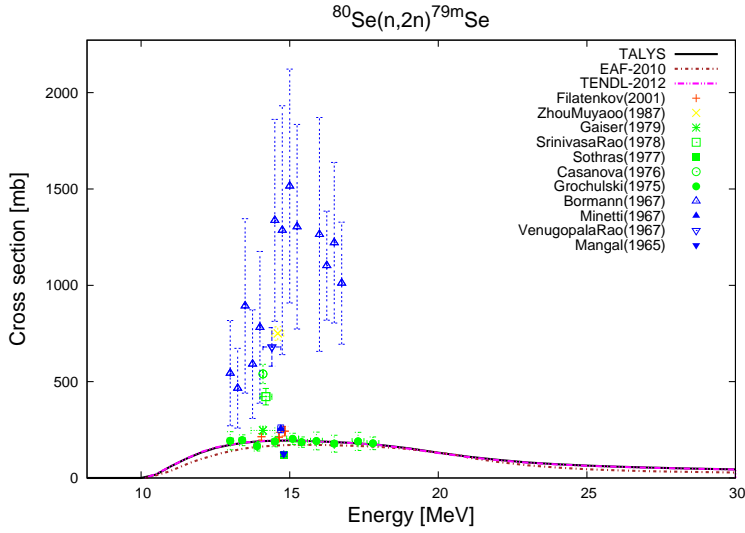
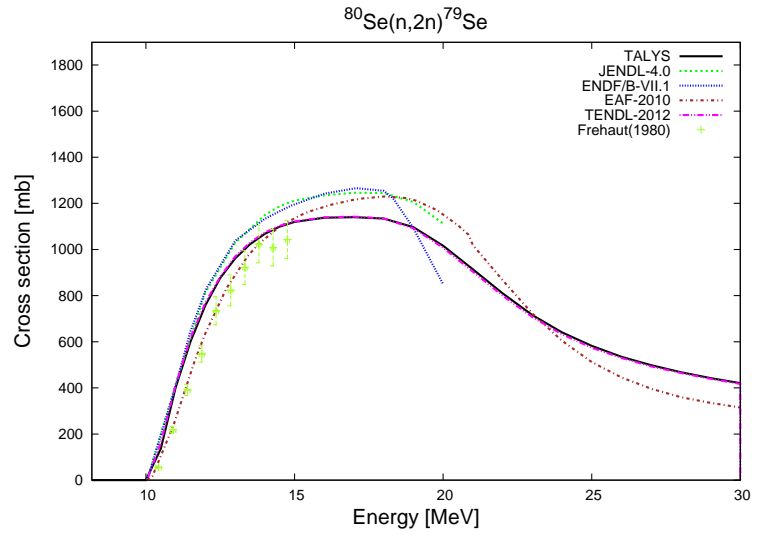
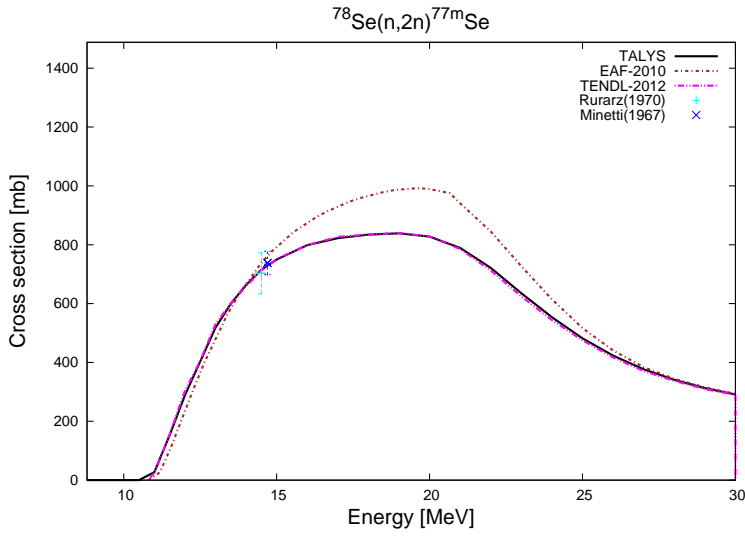


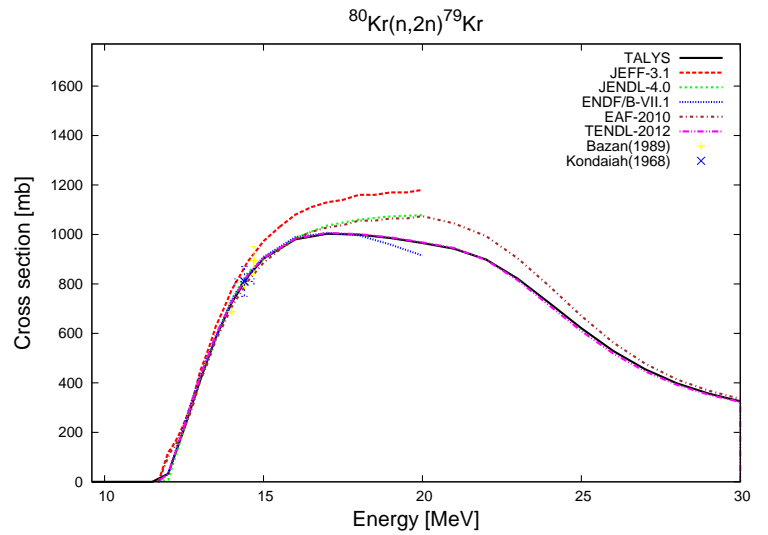
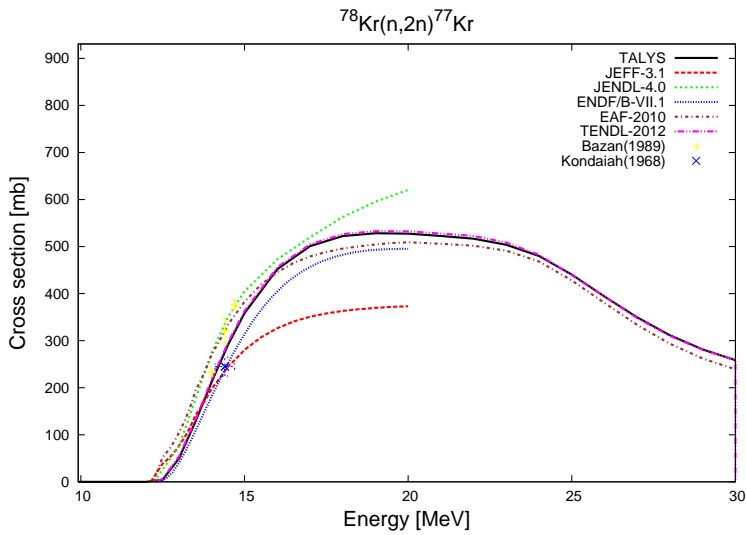
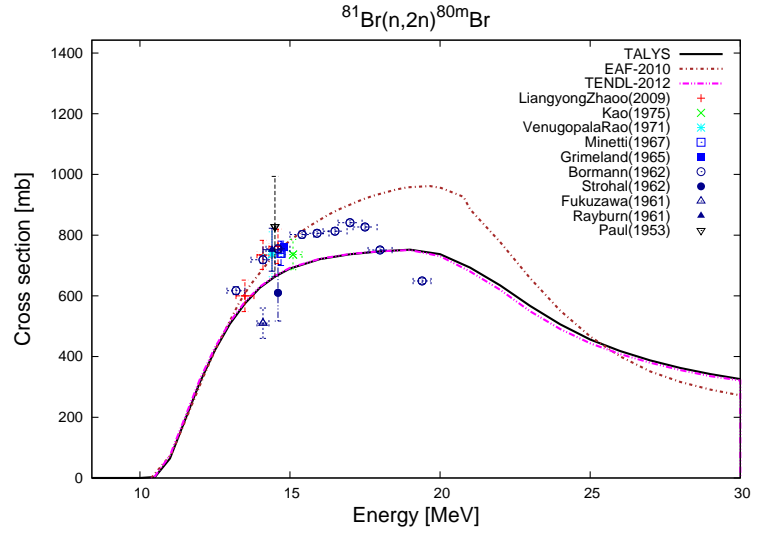
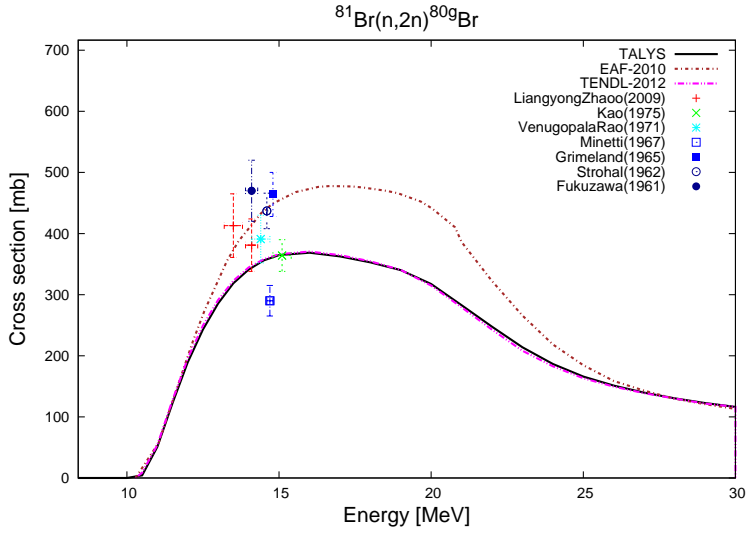
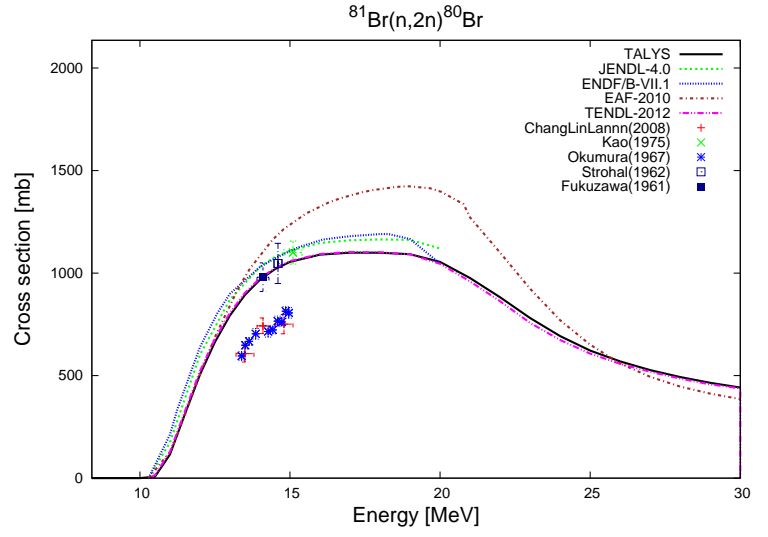
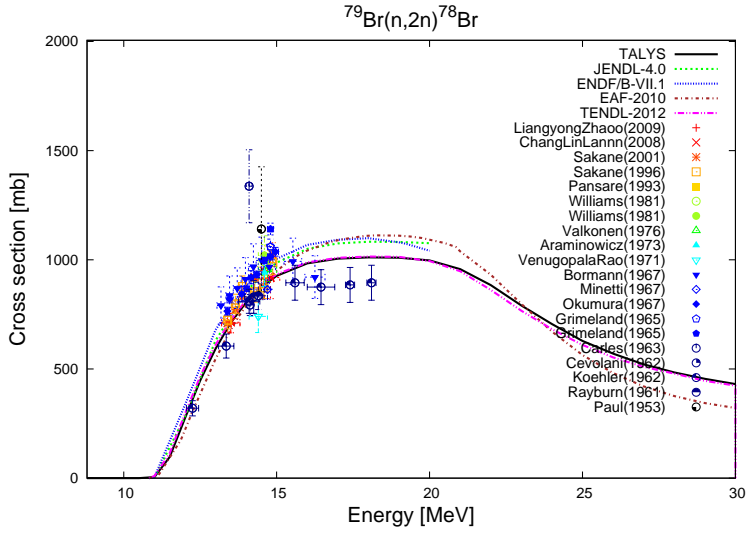


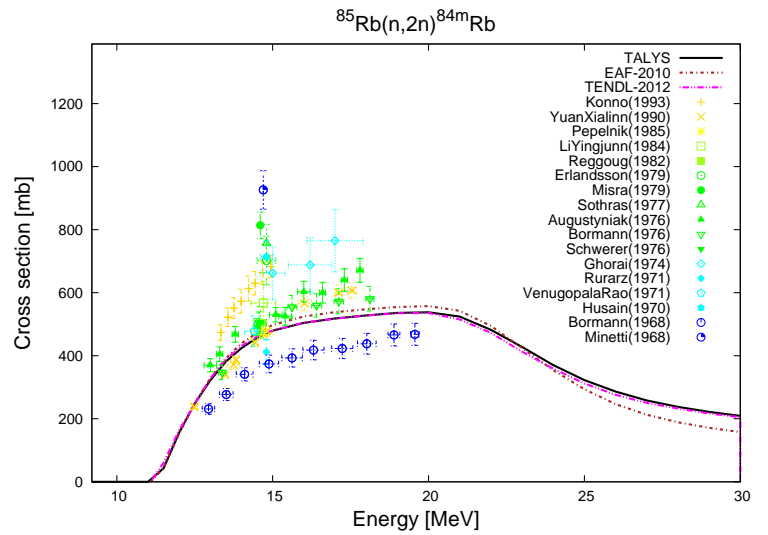
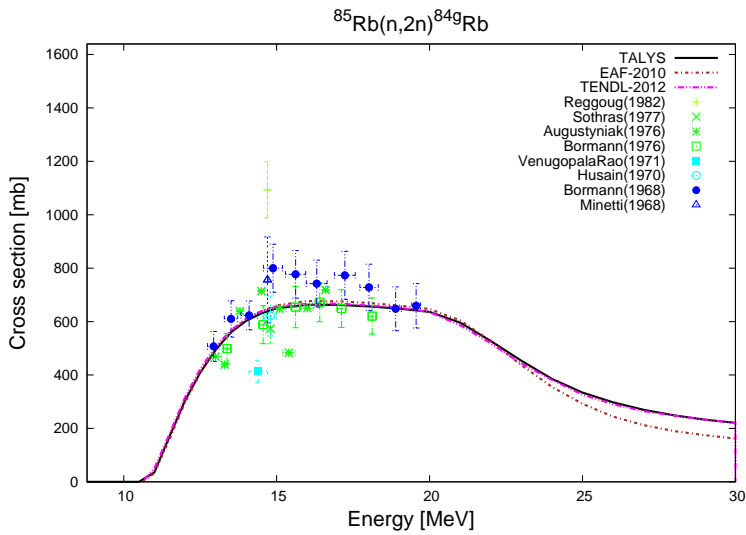
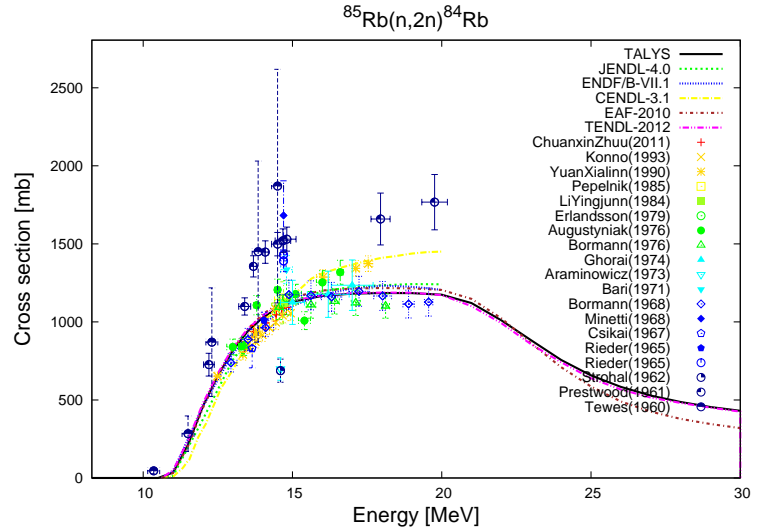
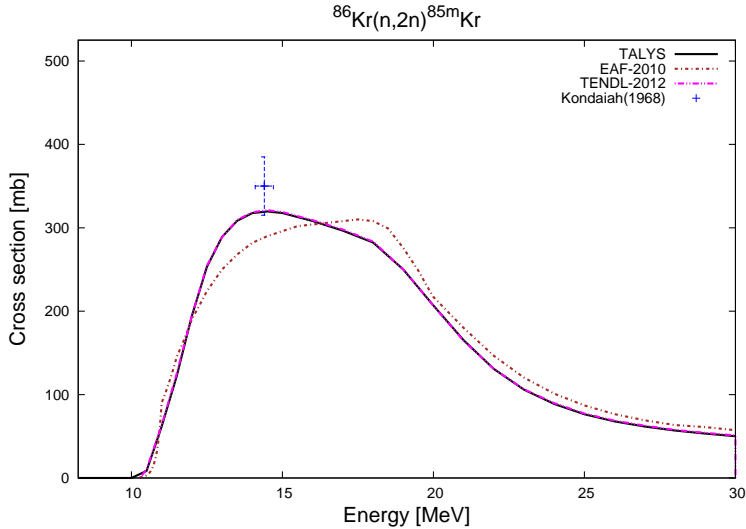
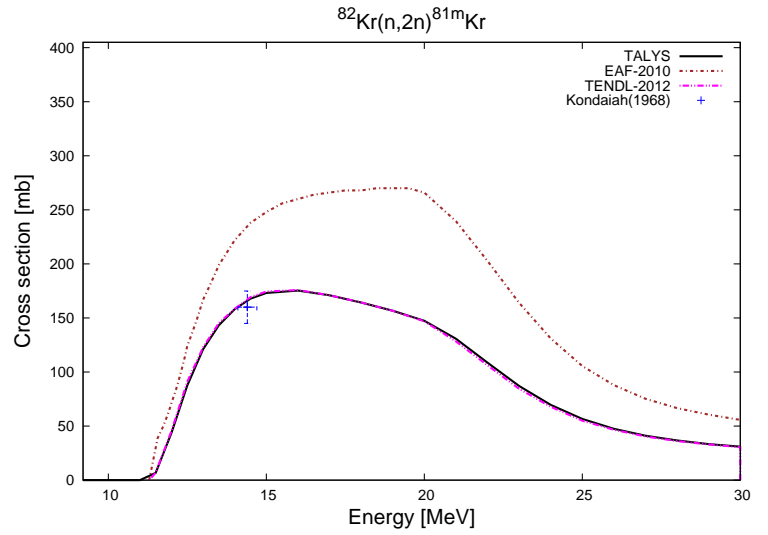
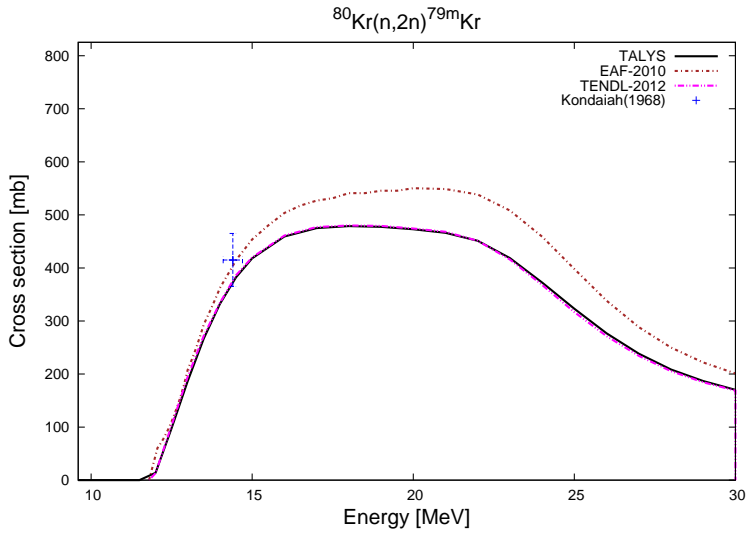


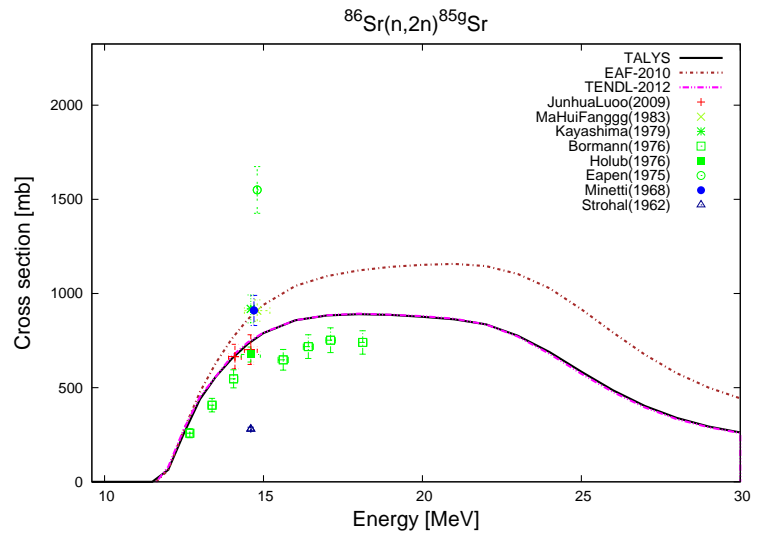
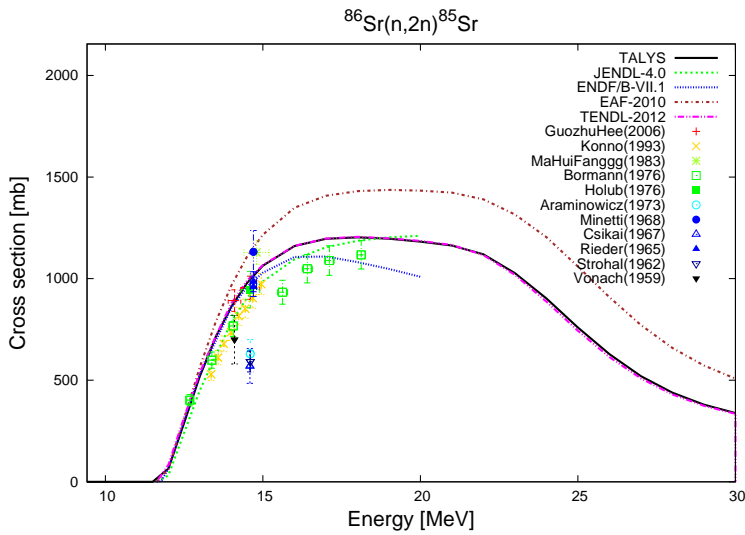
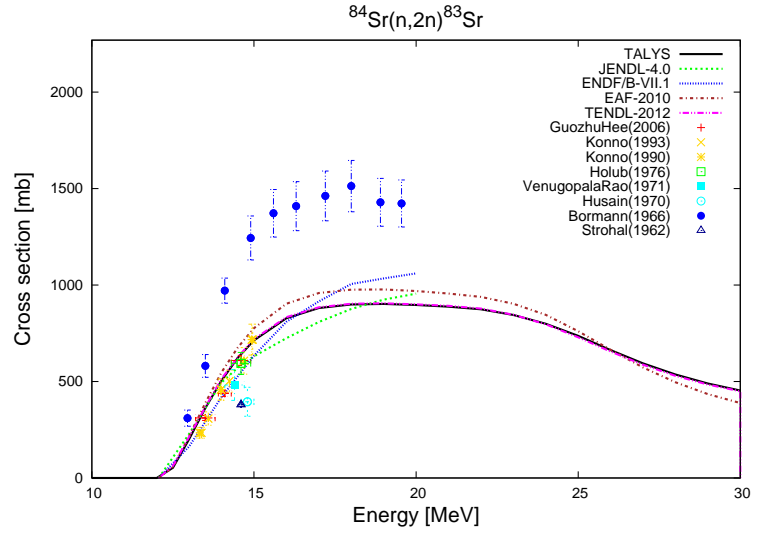
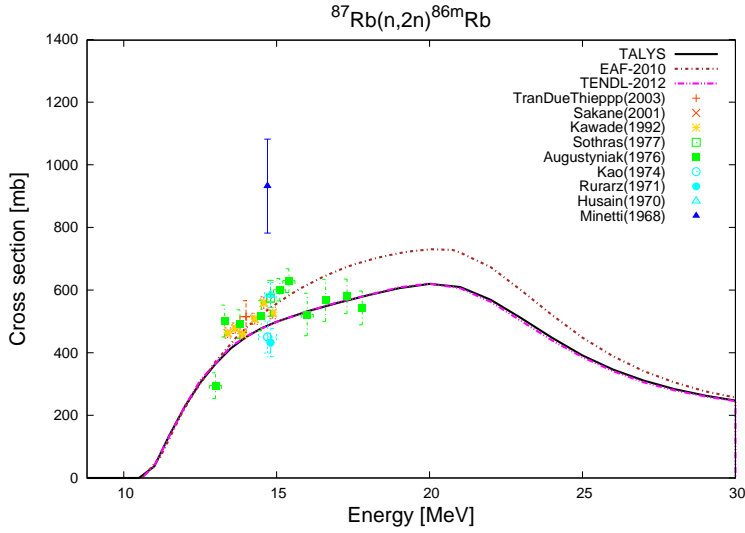
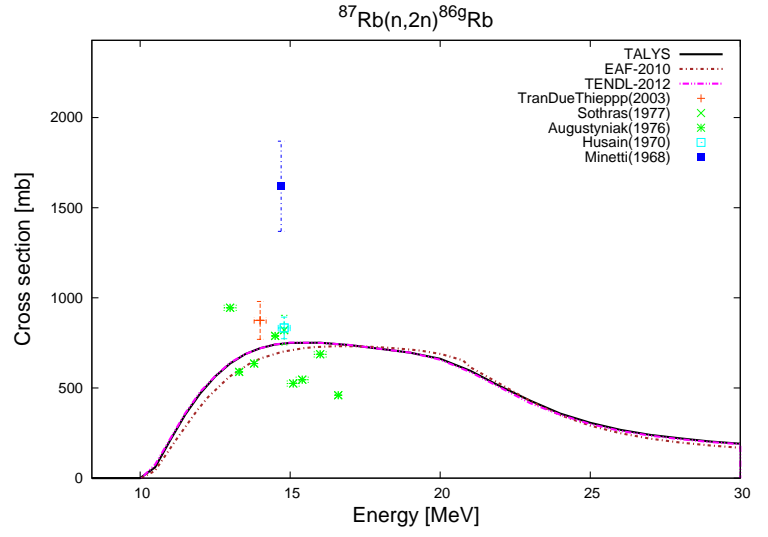
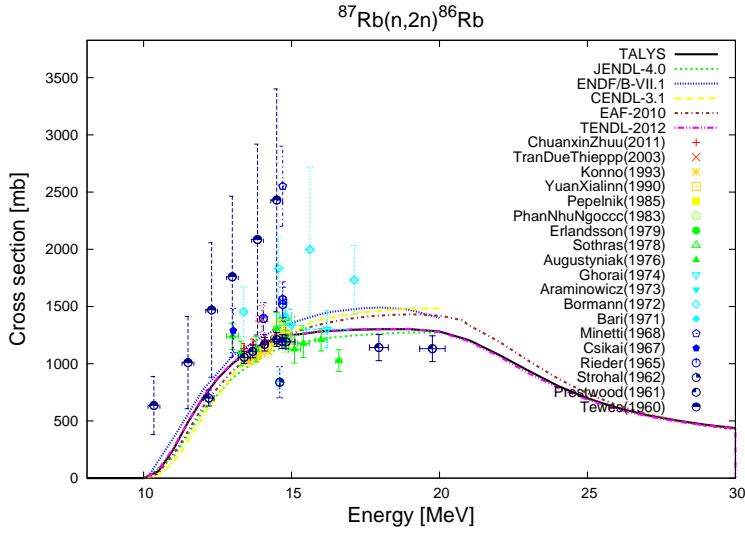
$^{71}\text{Ga}(n,2n)^{70}\text{Ga}$  $^{70}\text{Ge}(n,2n)^{69}\text{Ge}$  $^{72}\text{Ge}(n,2n)^{71}\text{Ge}$  $^{76}\text{Ge}(n,2n)^{75}\text{Ge}$  $^{76}\text{Ge}(n,2n)^{75g}\text{Ge}$  $^{76}\text{Ge}(n,2n)^{75m}\text{Ge}$ 

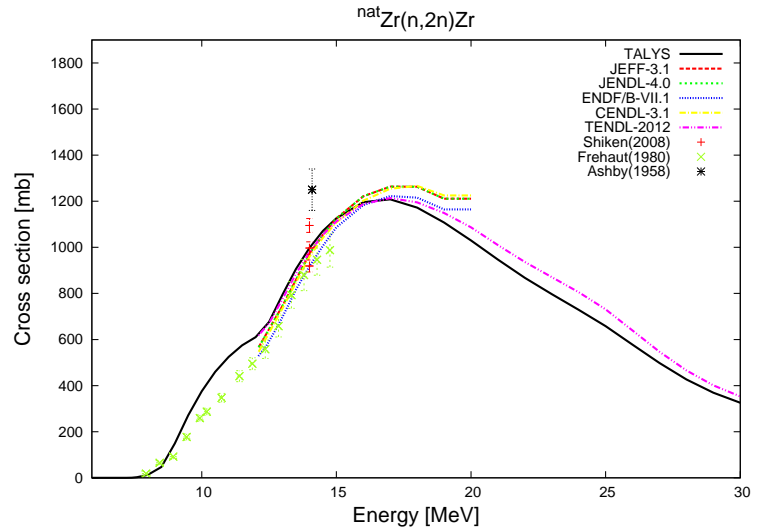
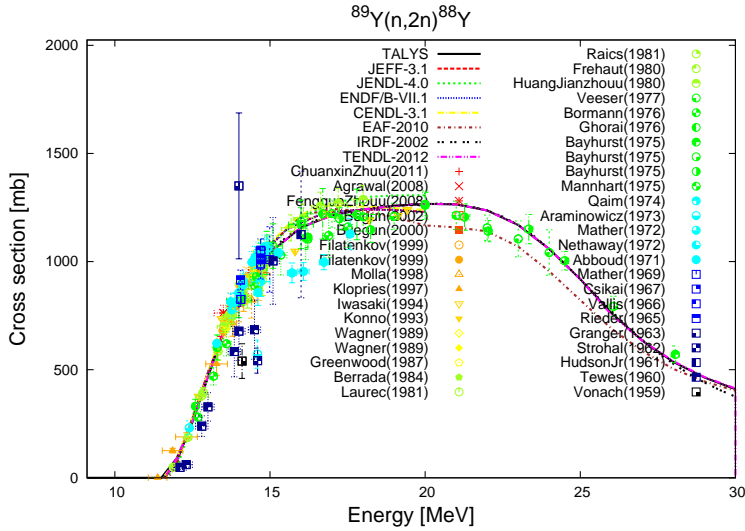
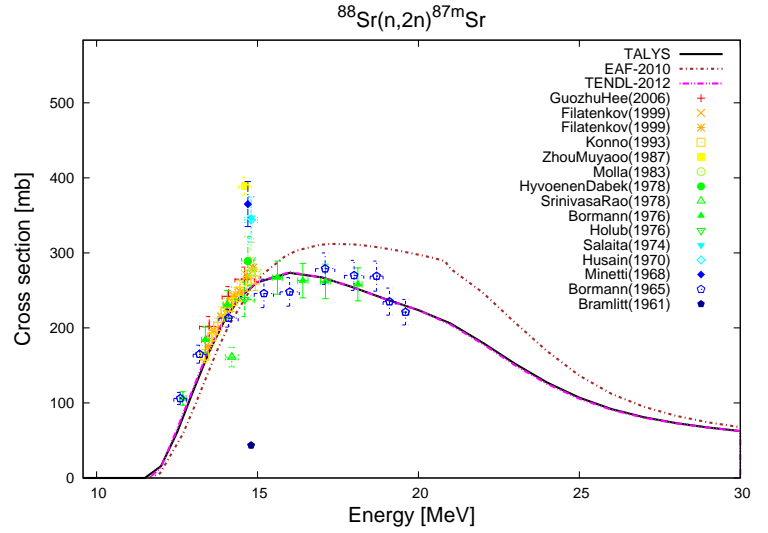
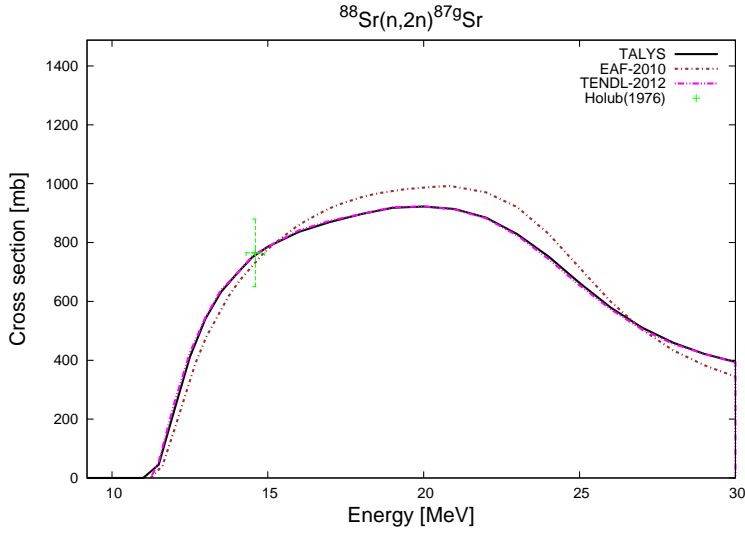
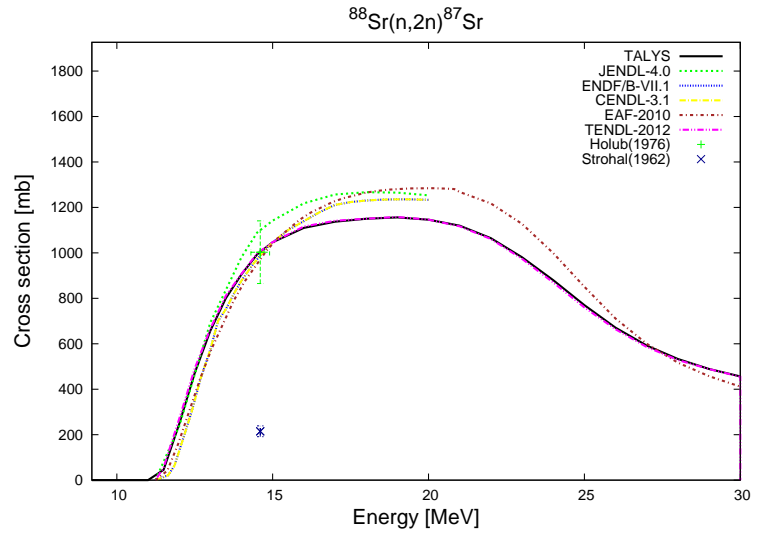
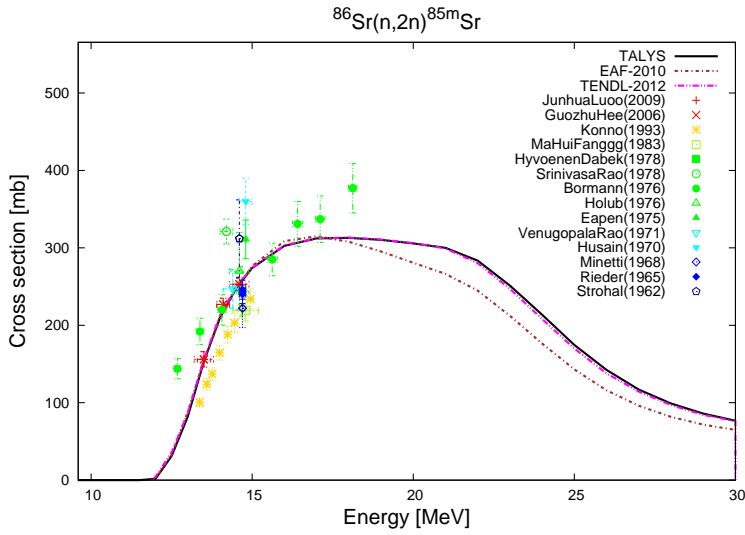


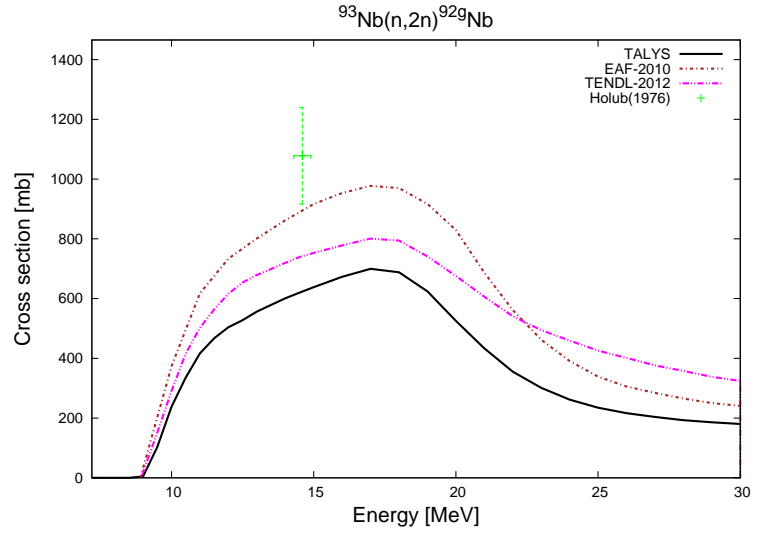
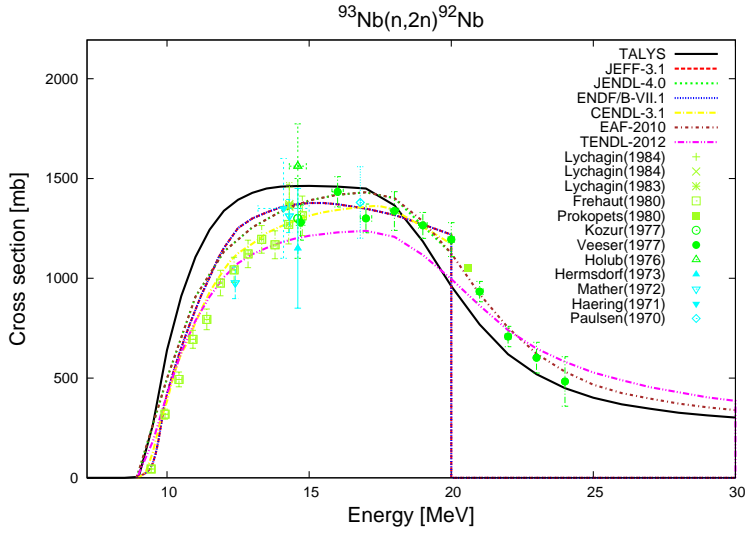
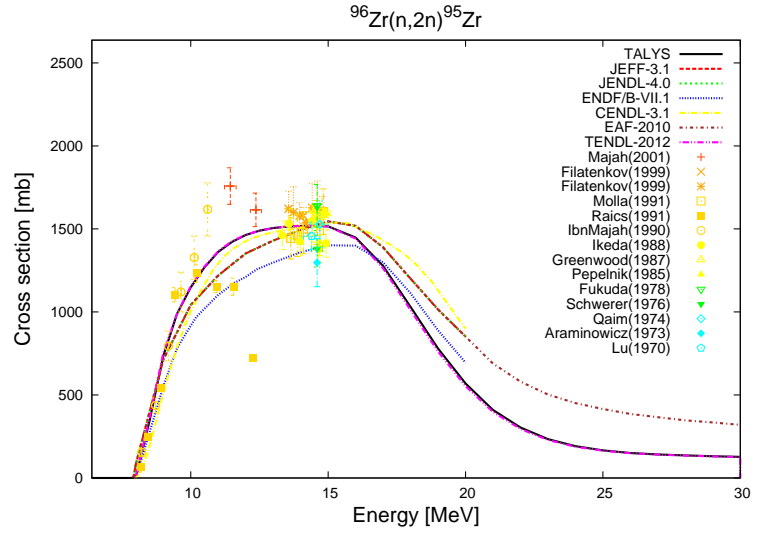
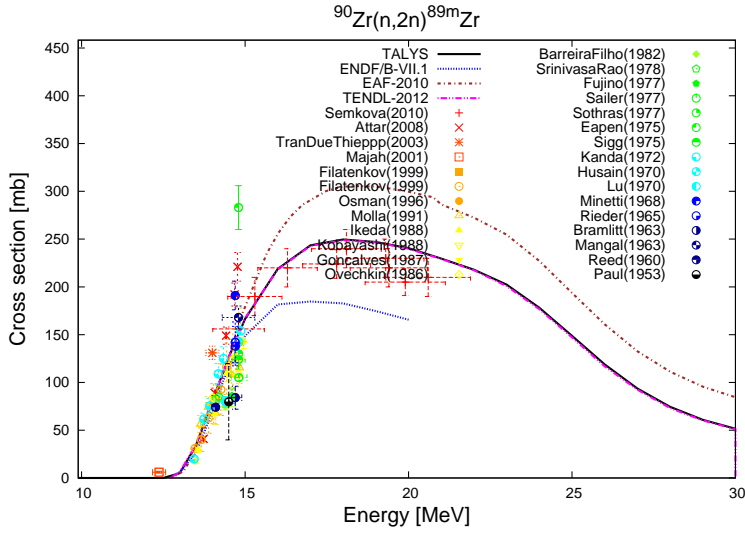
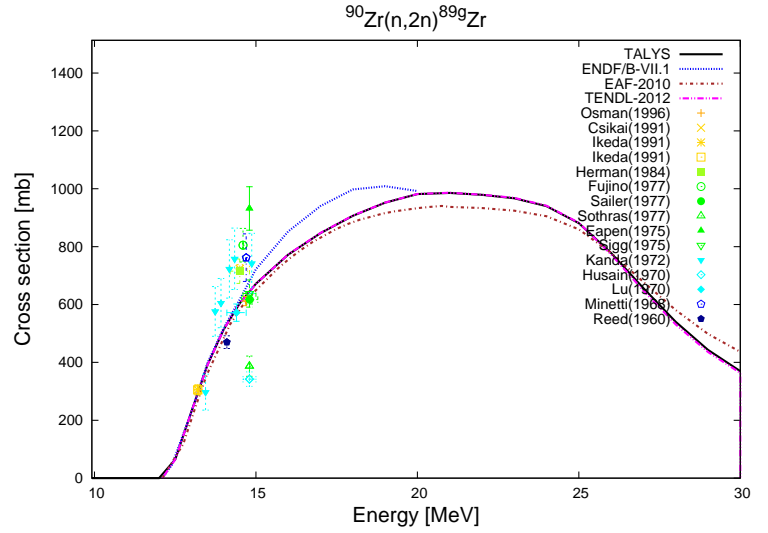
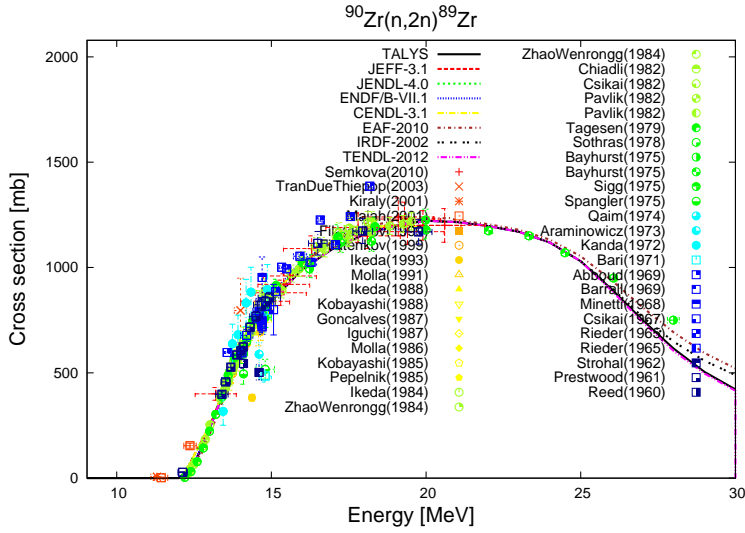


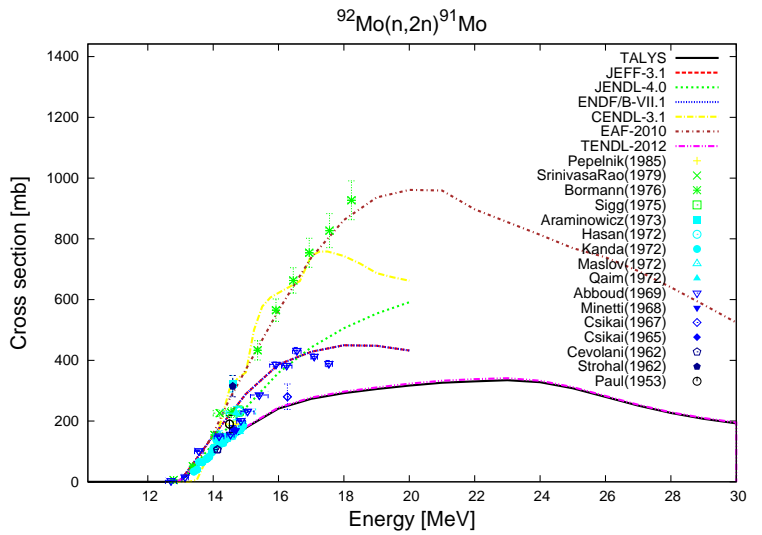
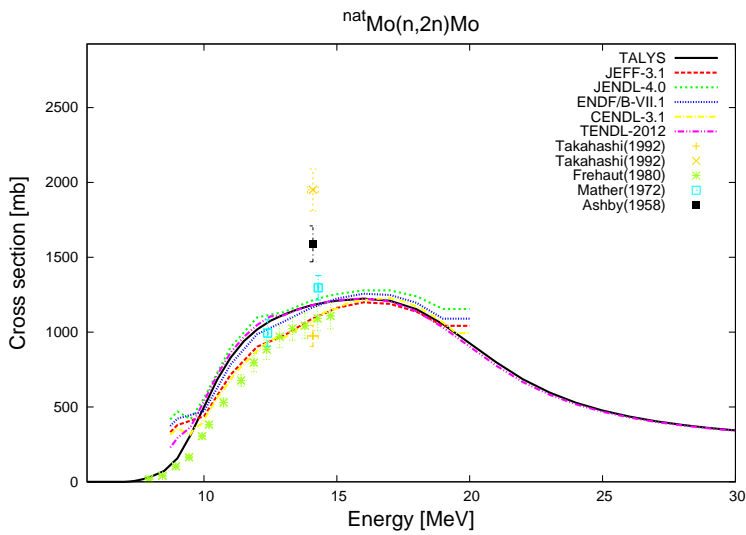
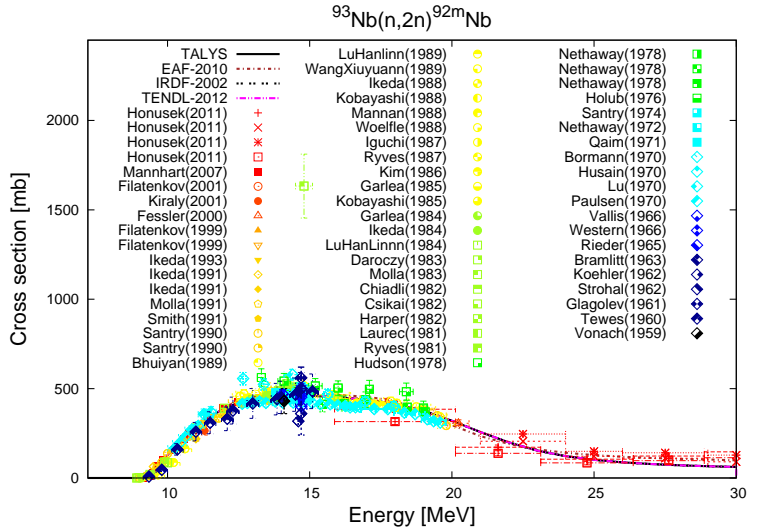
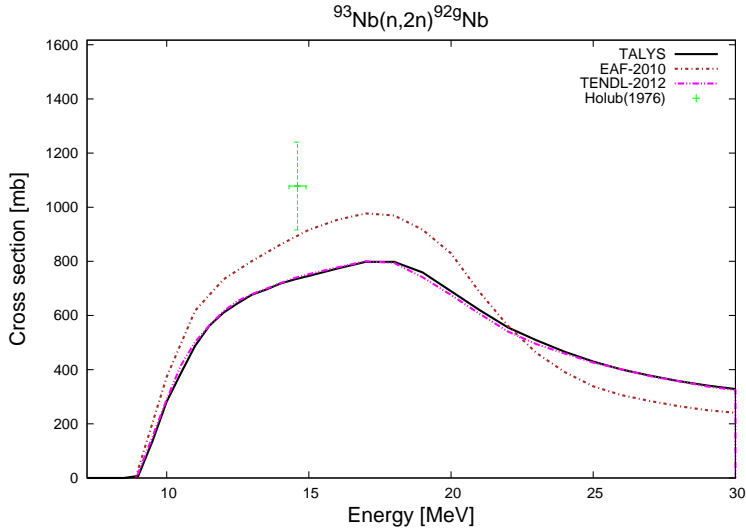
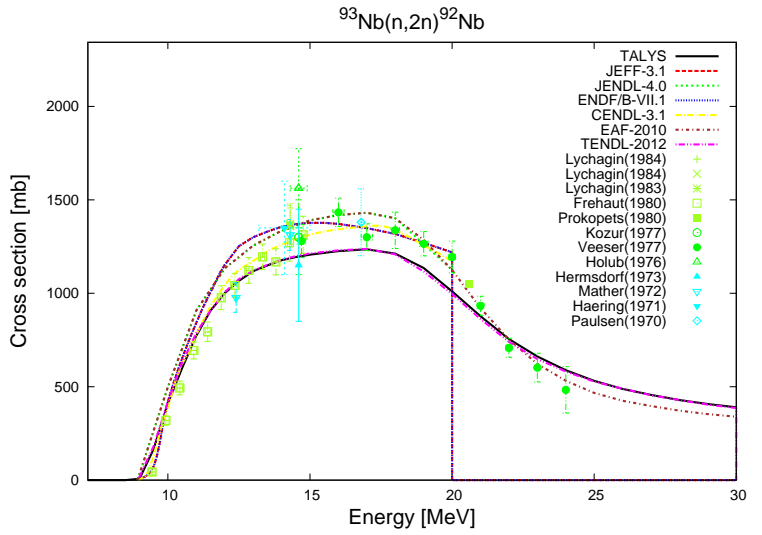
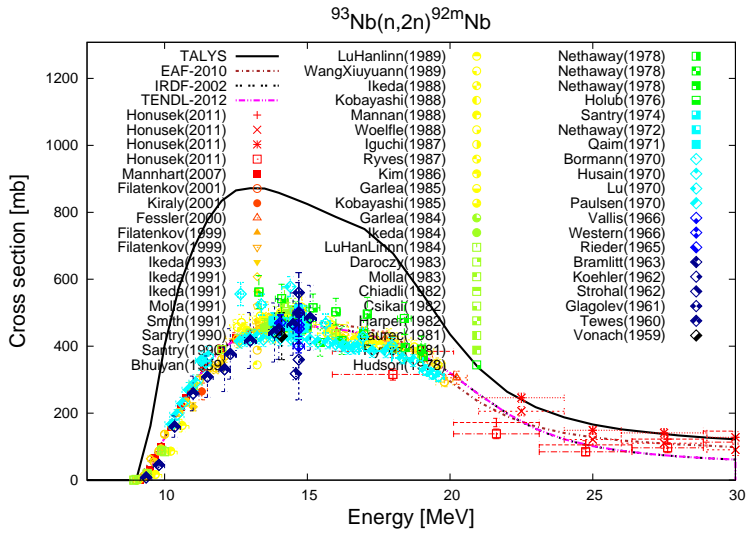


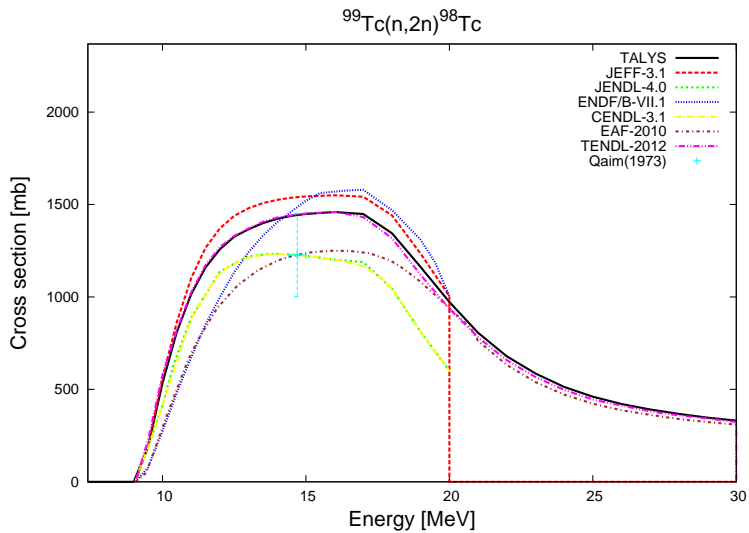
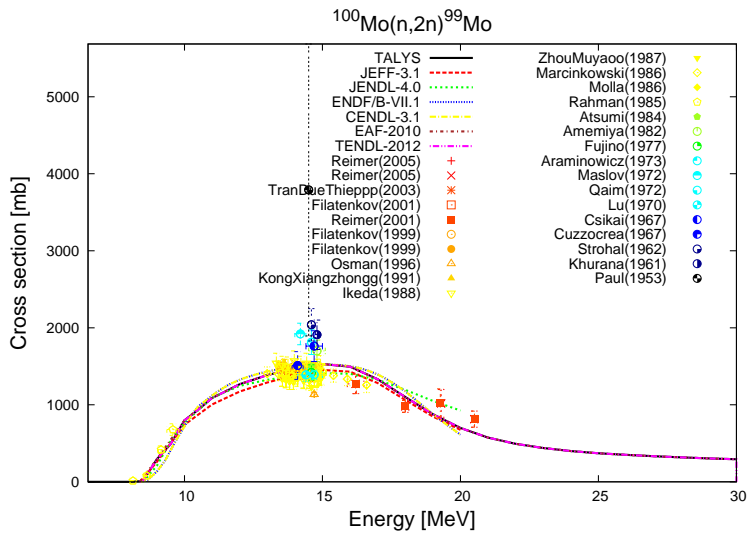
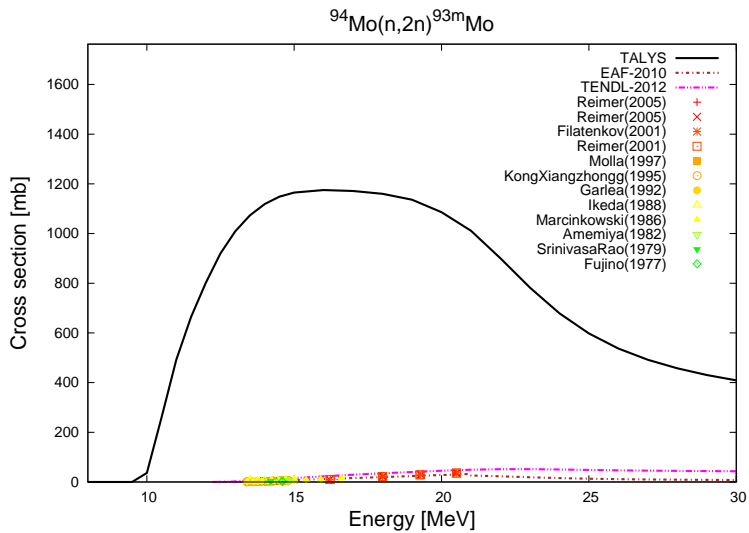
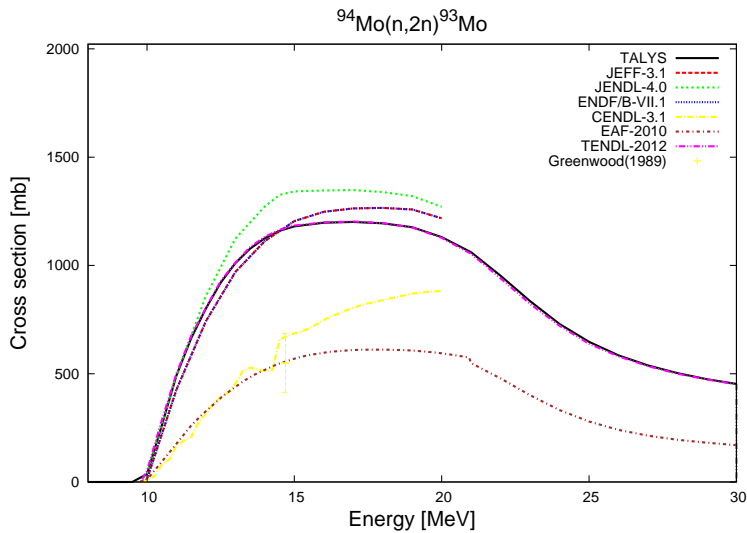
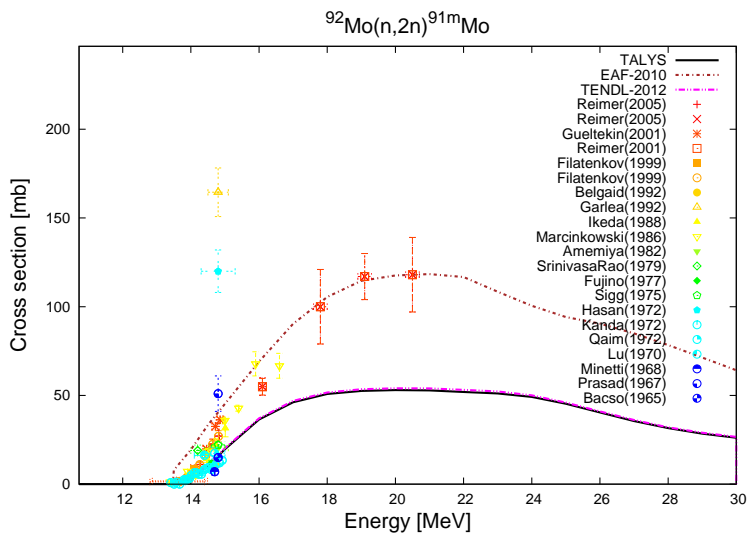
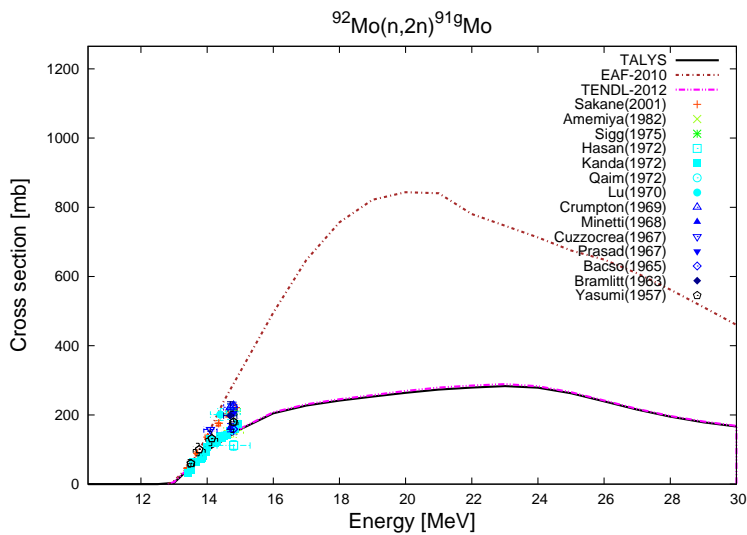


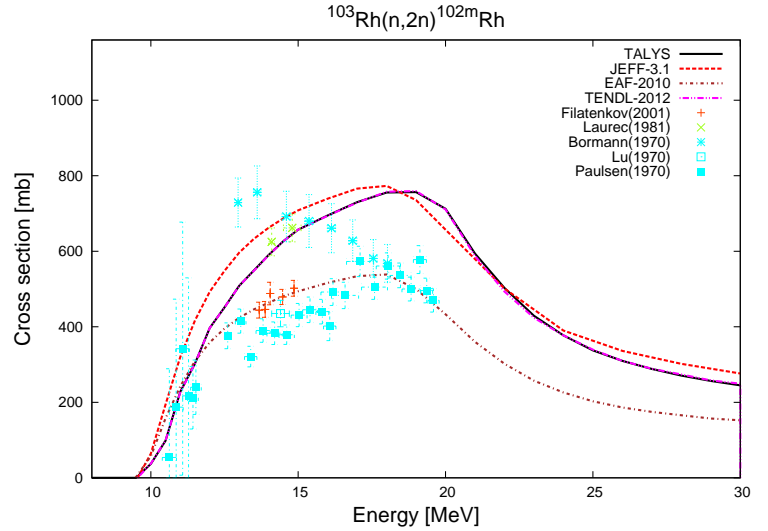
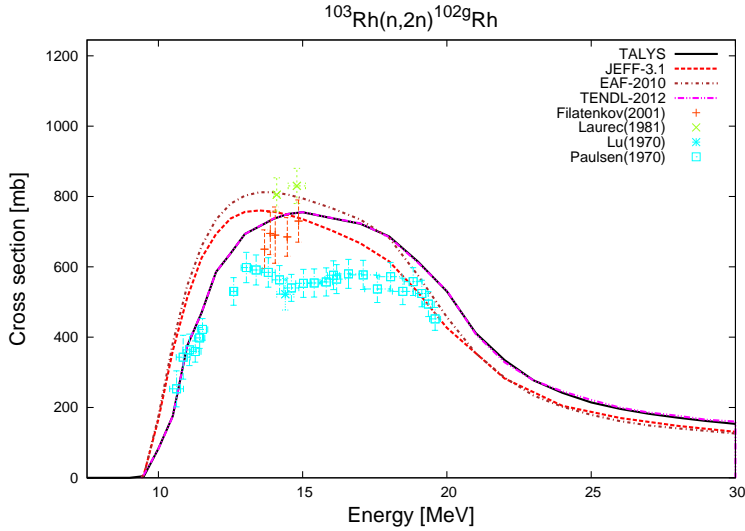
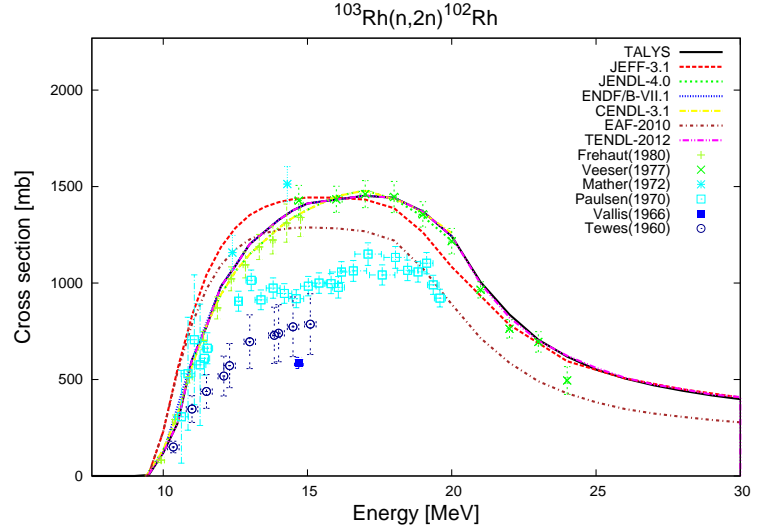
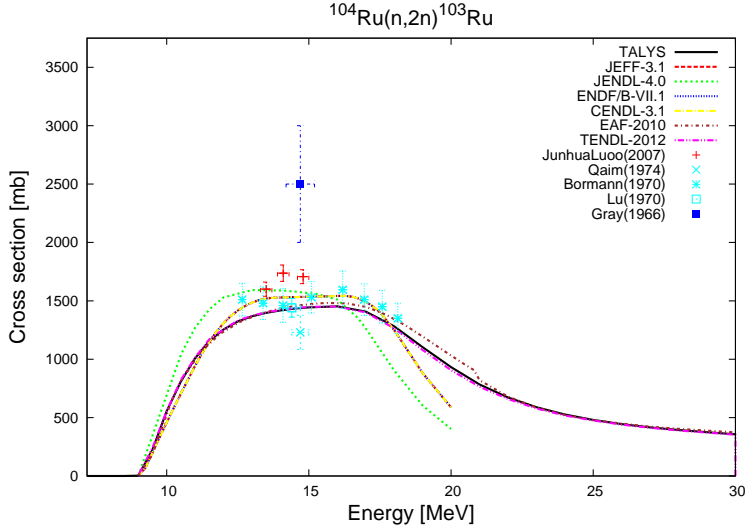
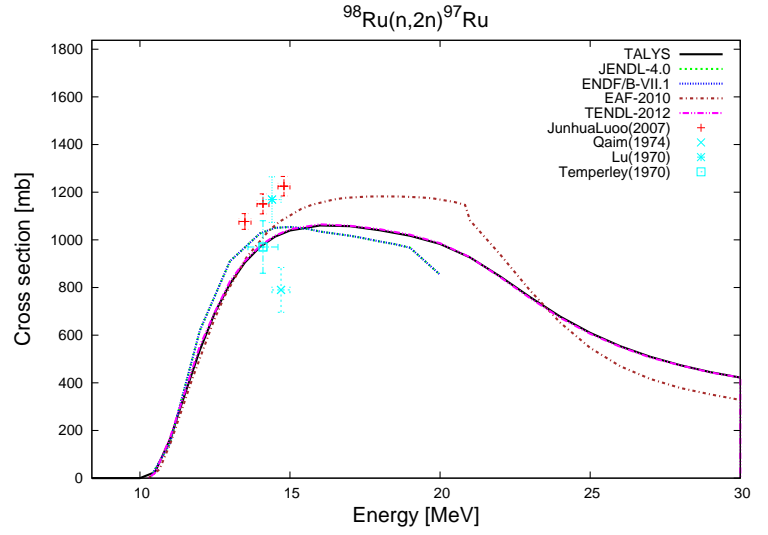
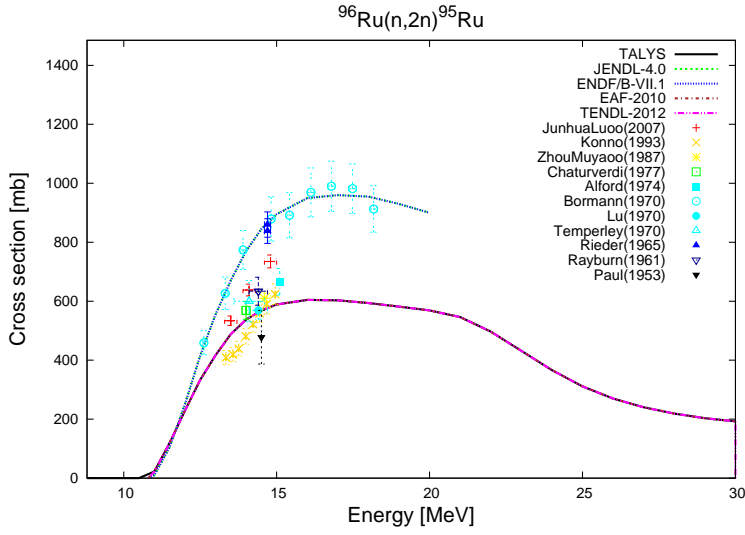


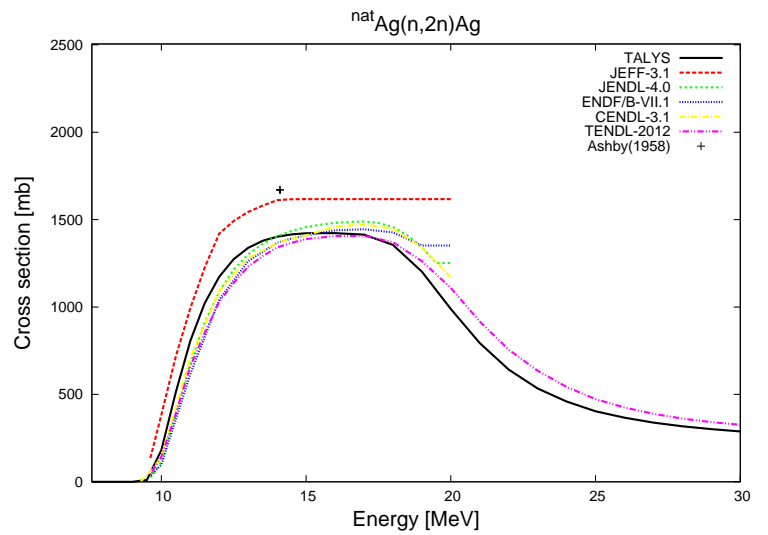
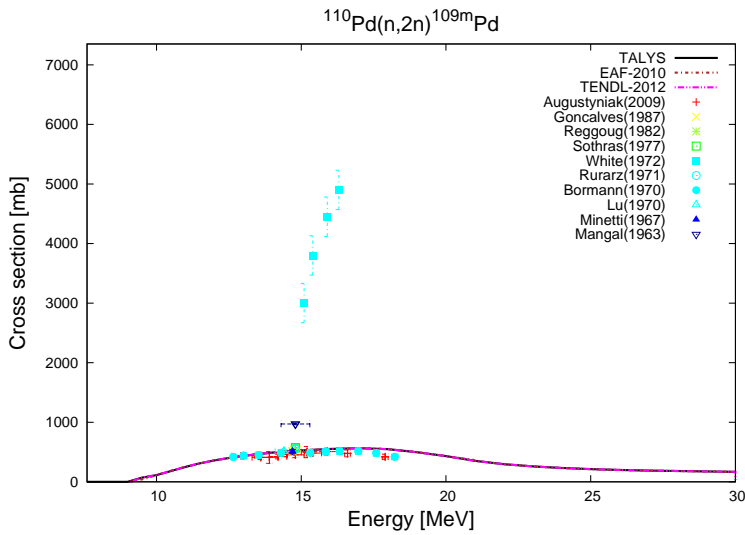
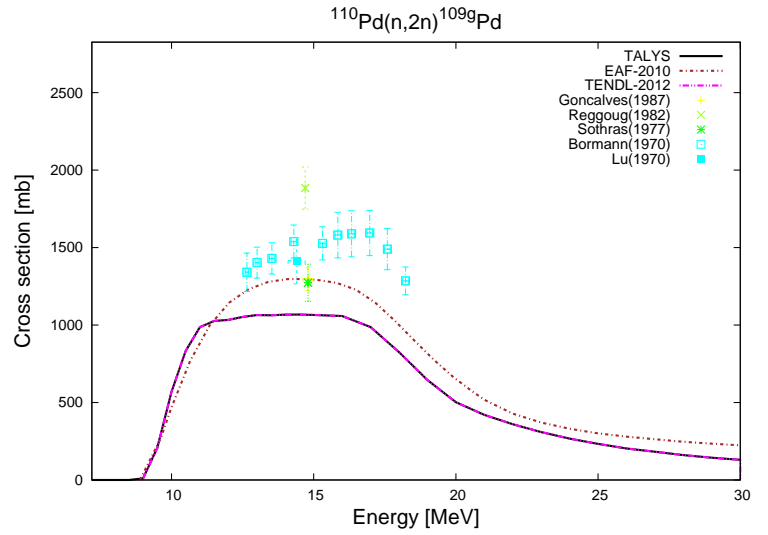
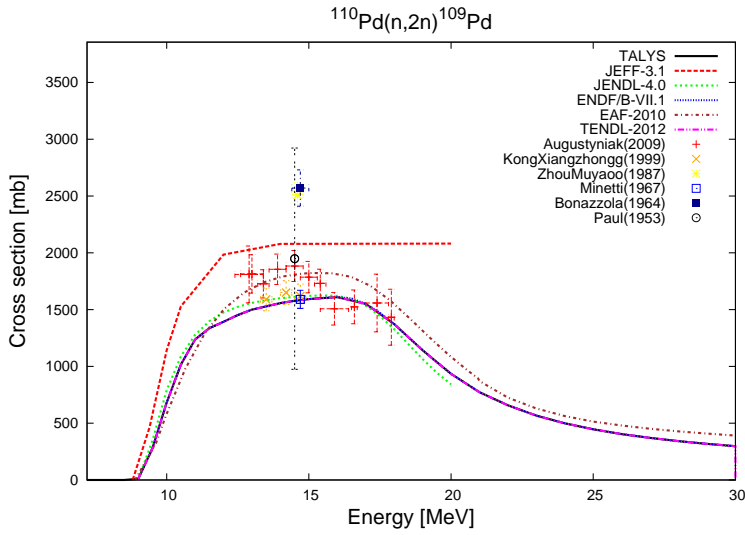
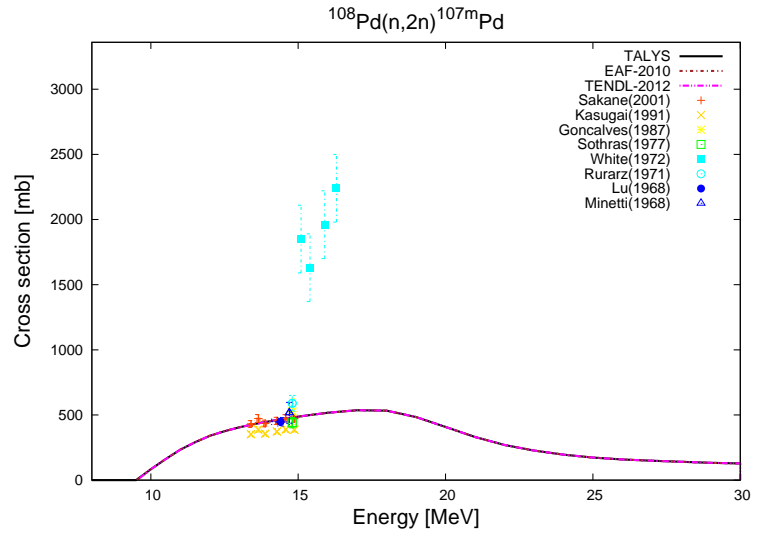
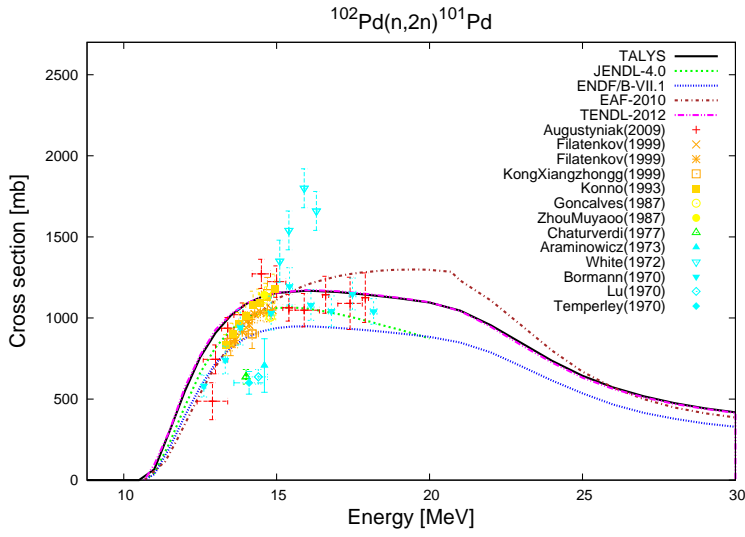


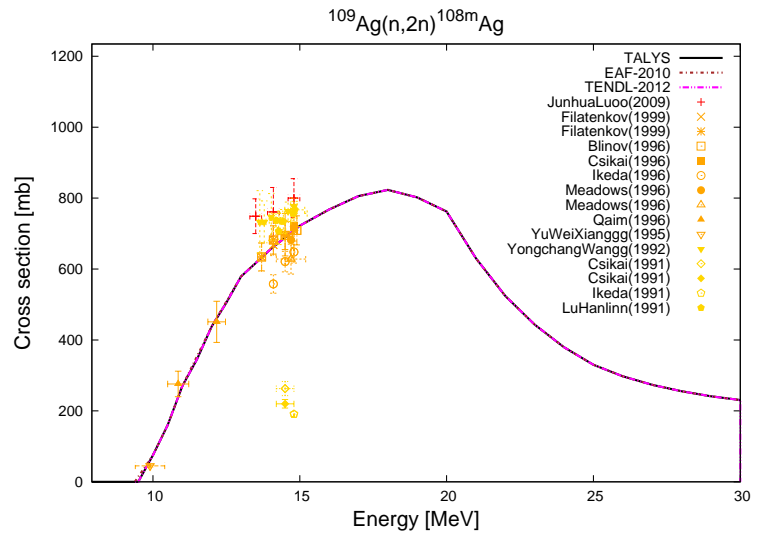
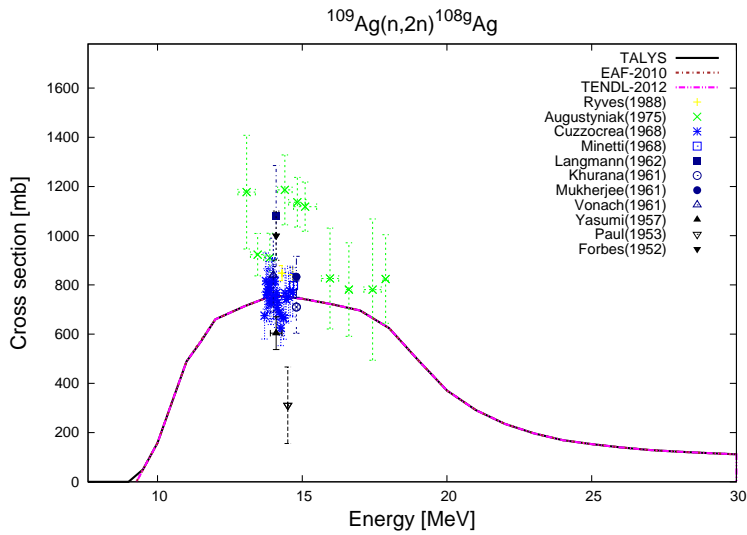
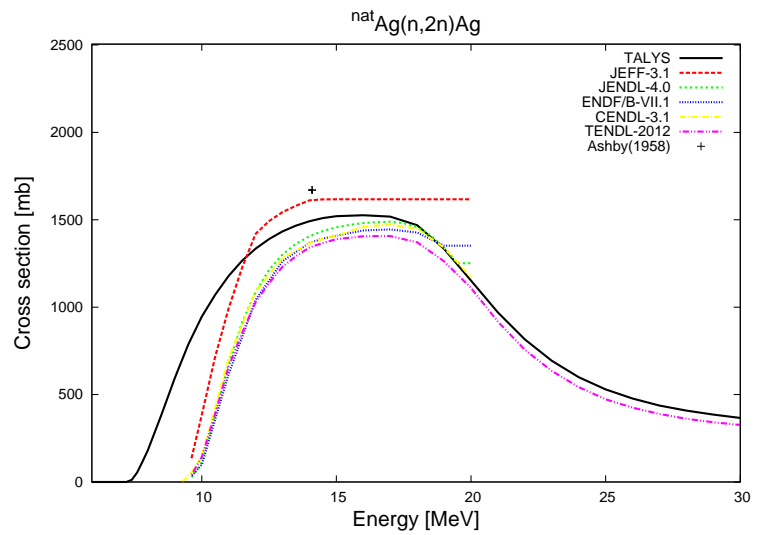
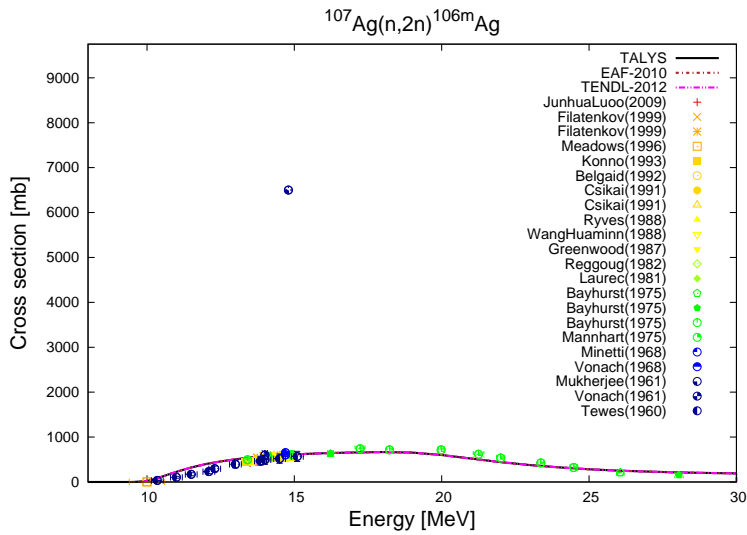
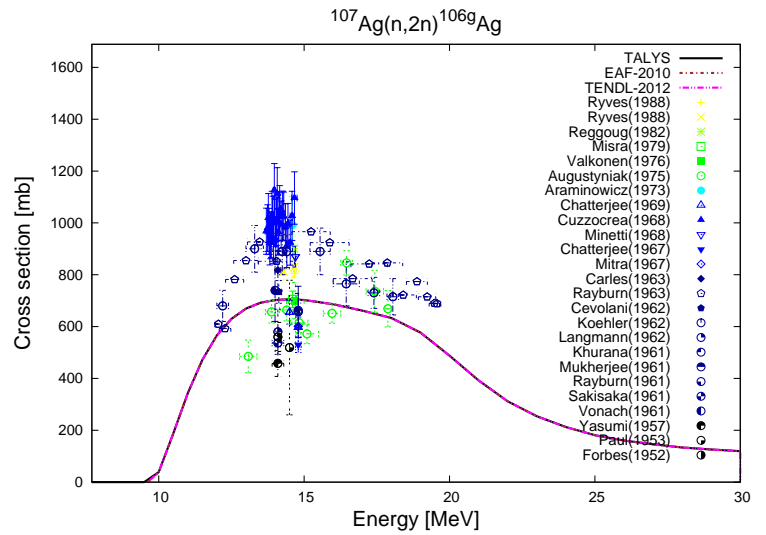
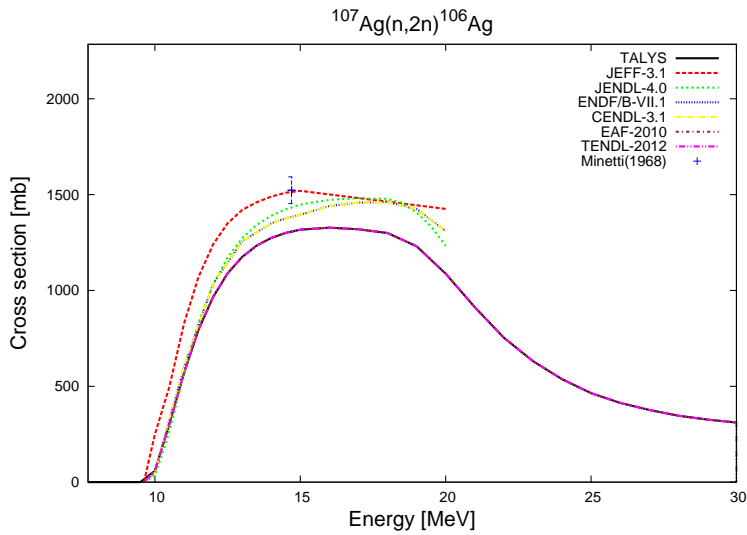


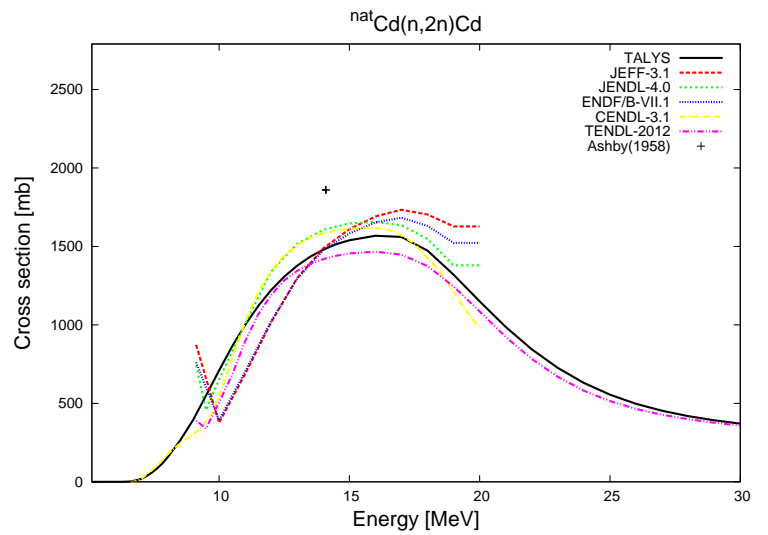
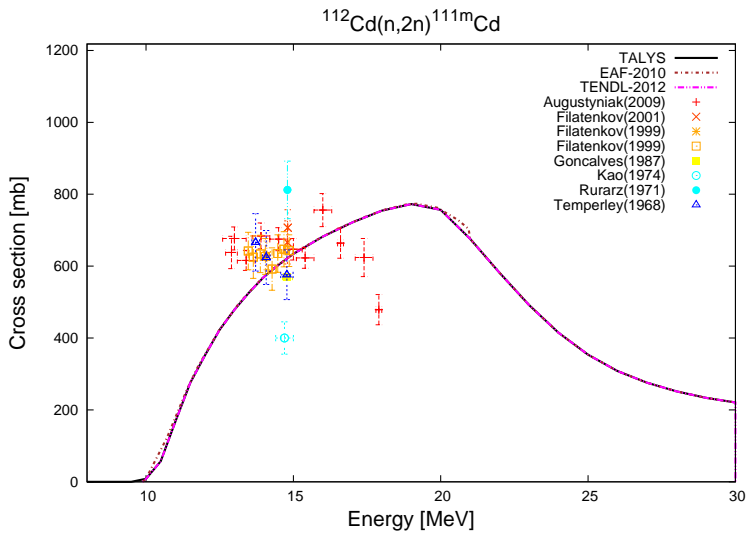
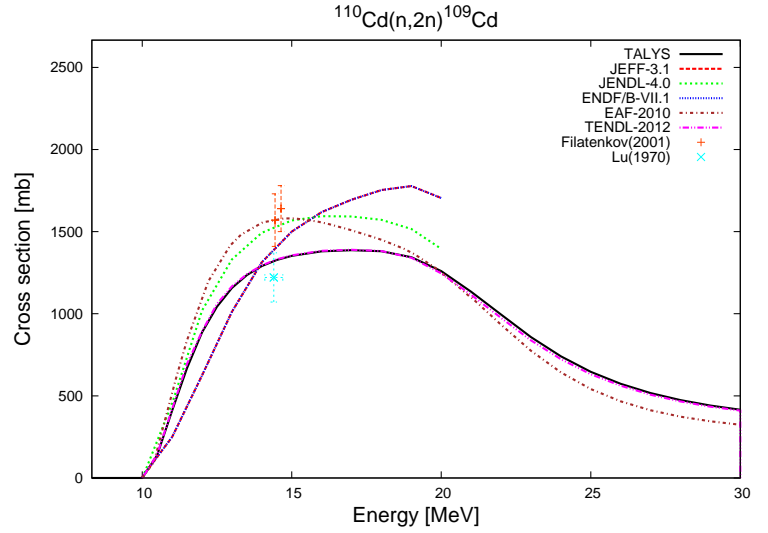
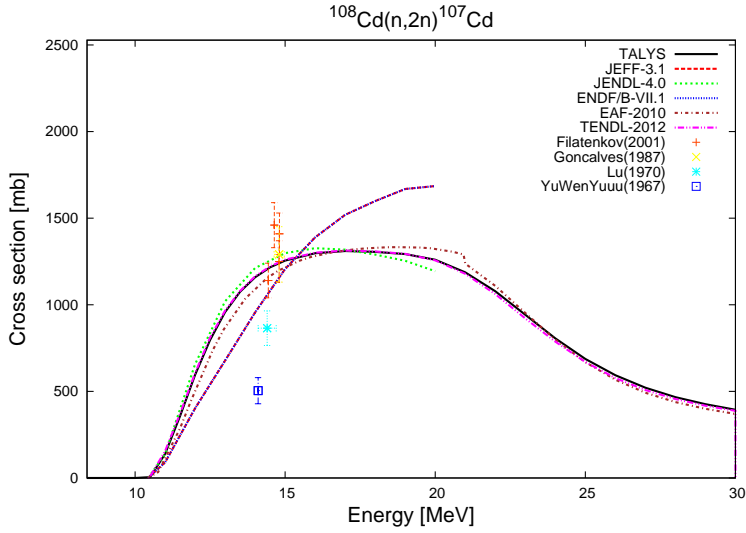
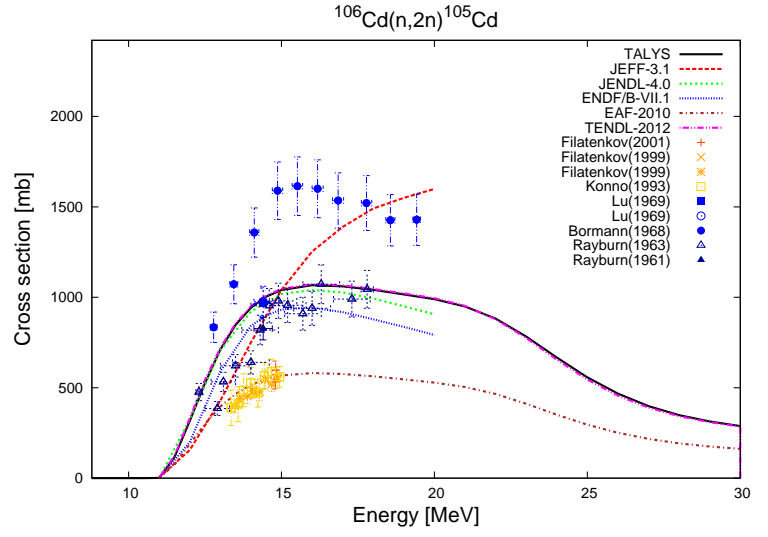
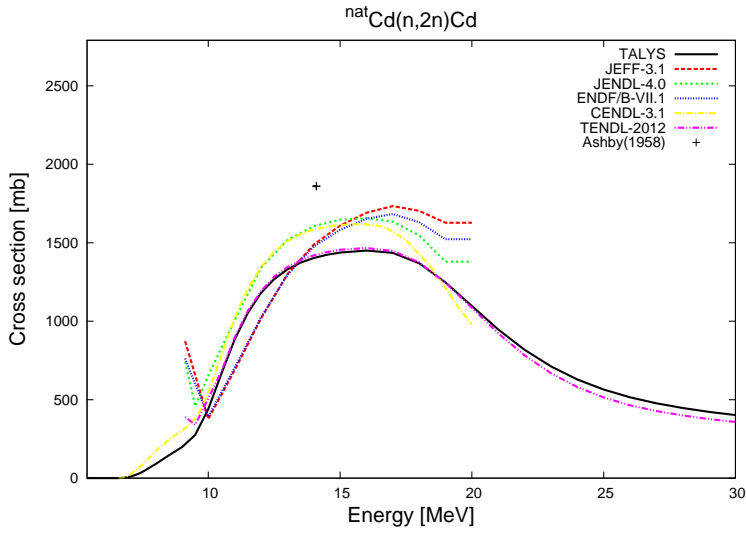


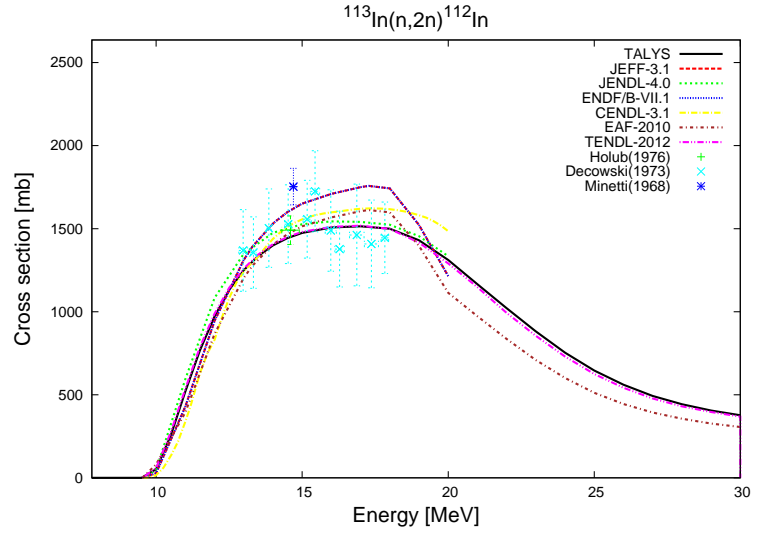
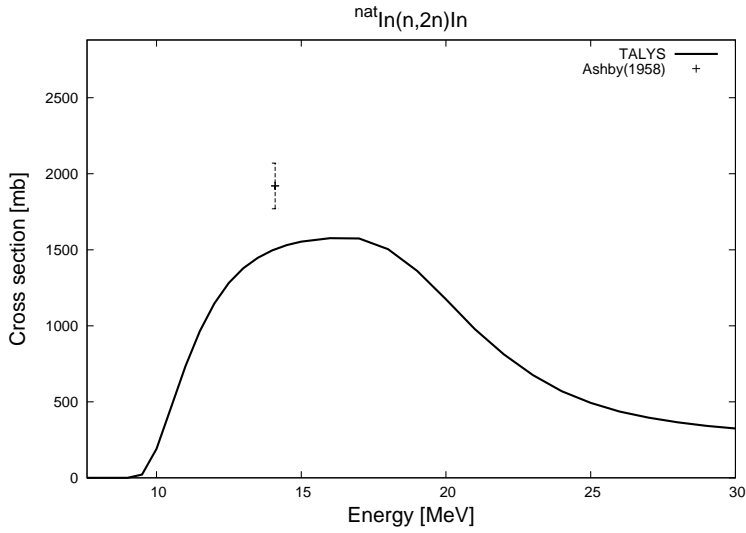
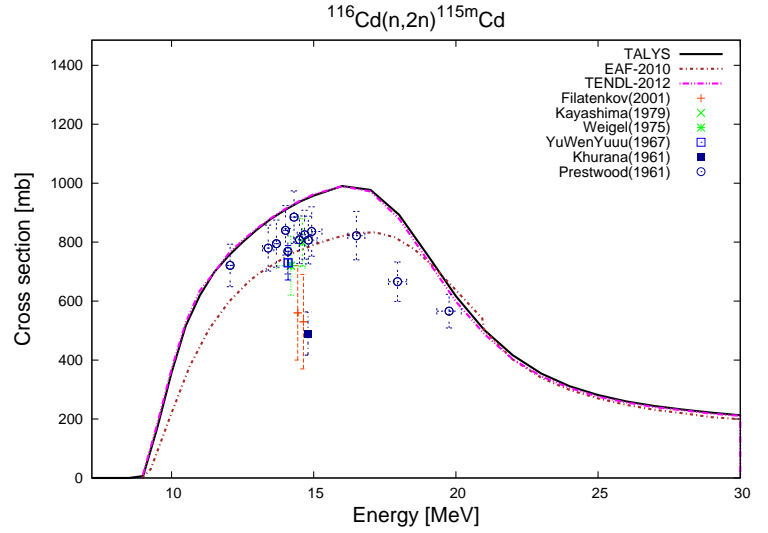
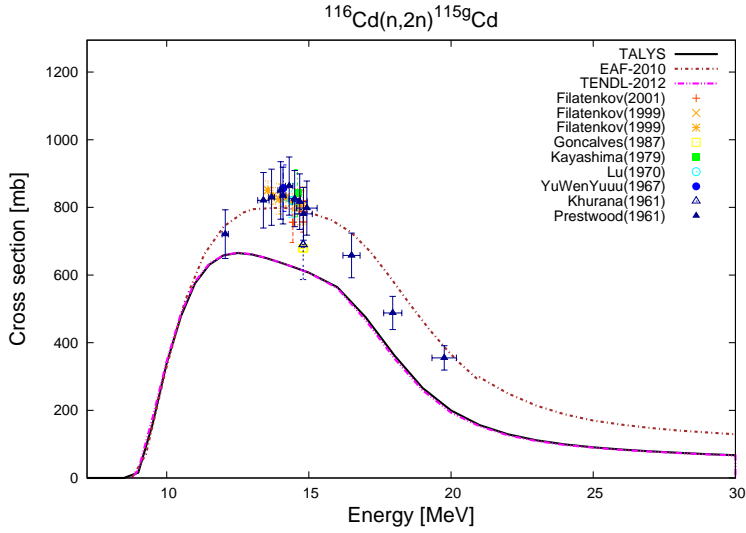
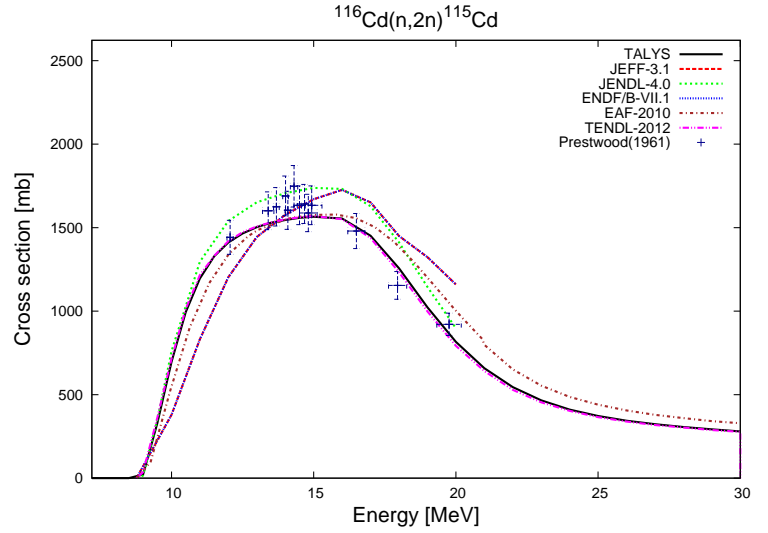
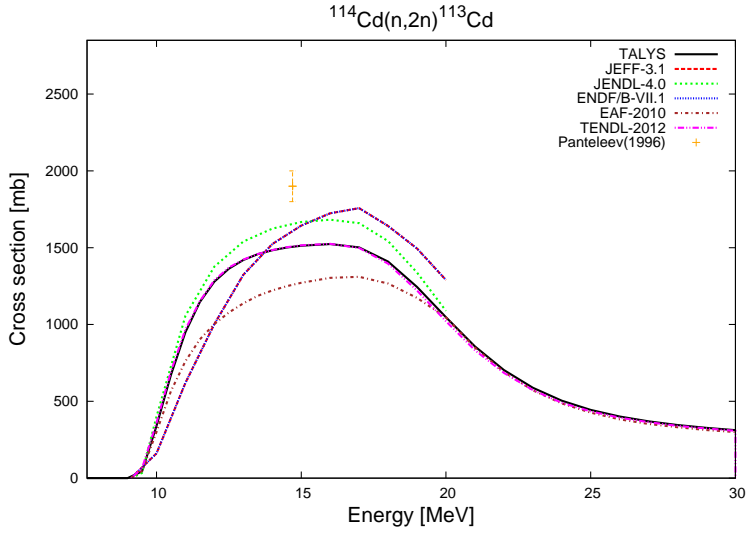


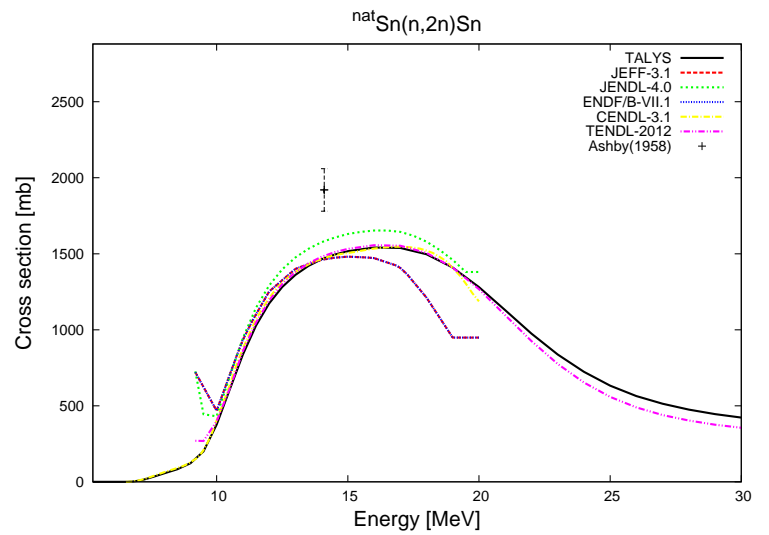
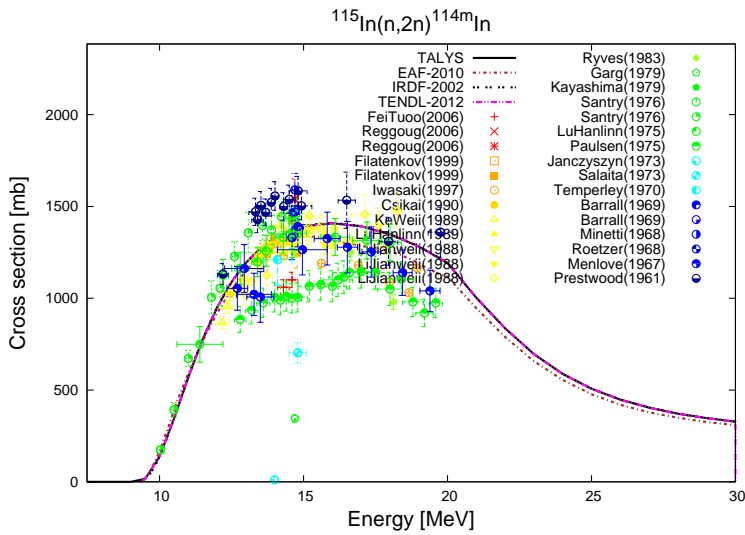
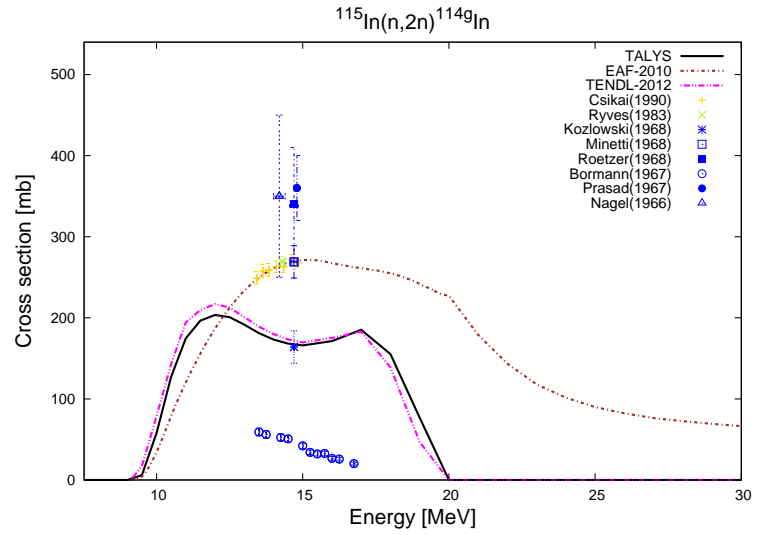
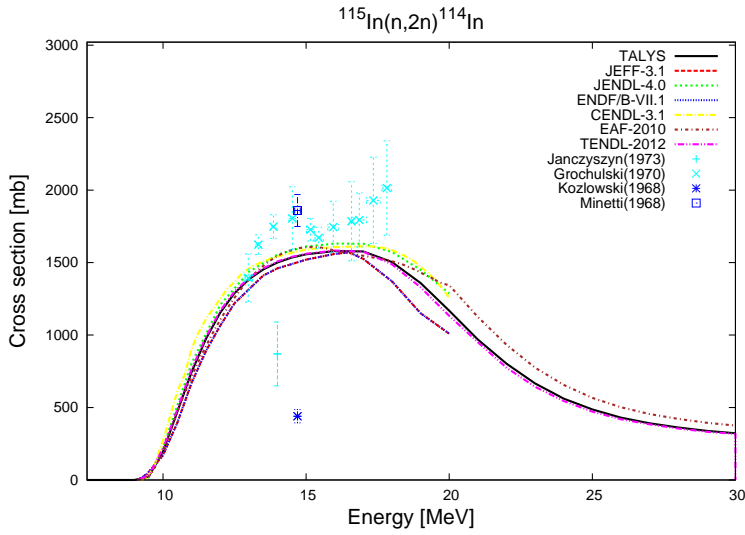
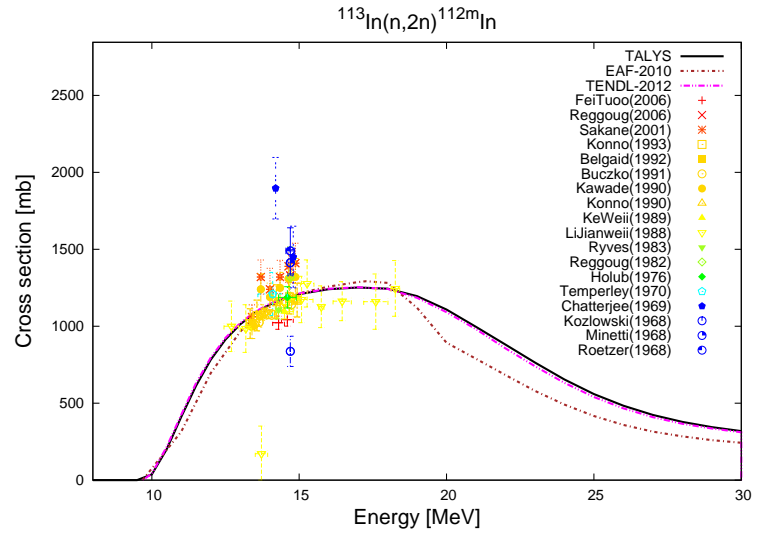
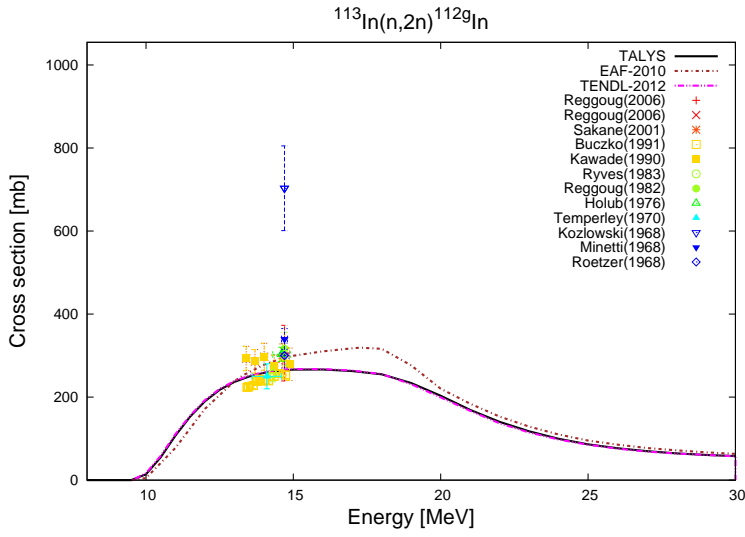


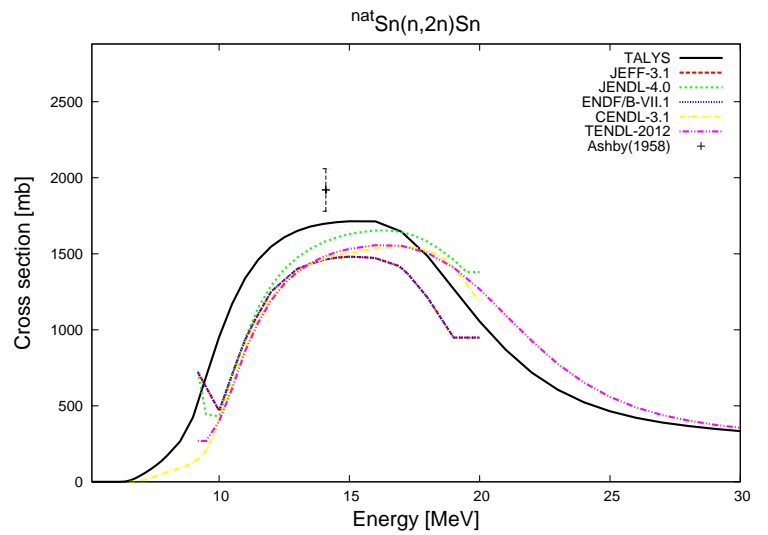
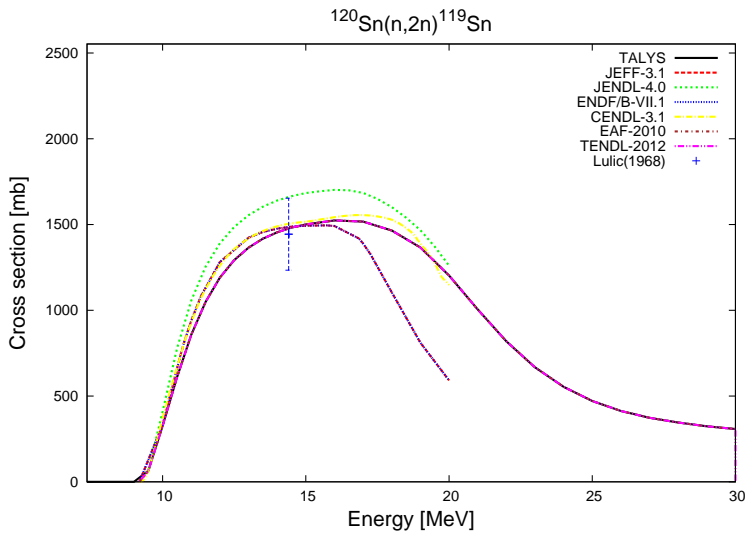
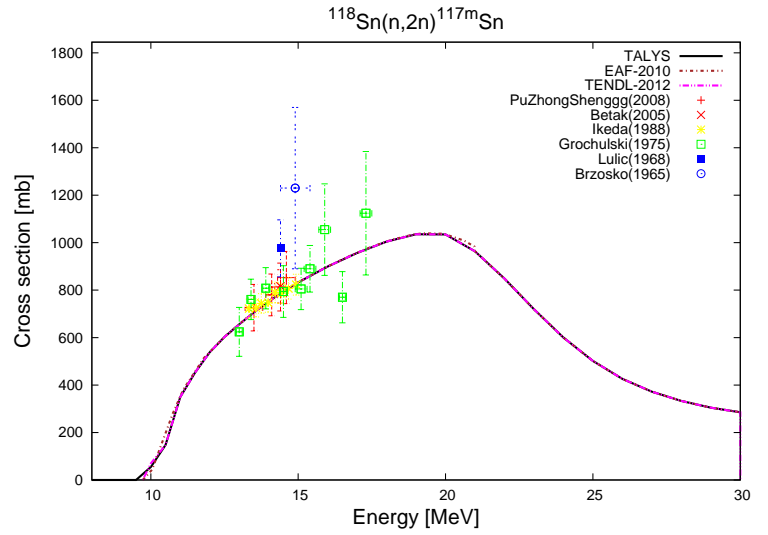
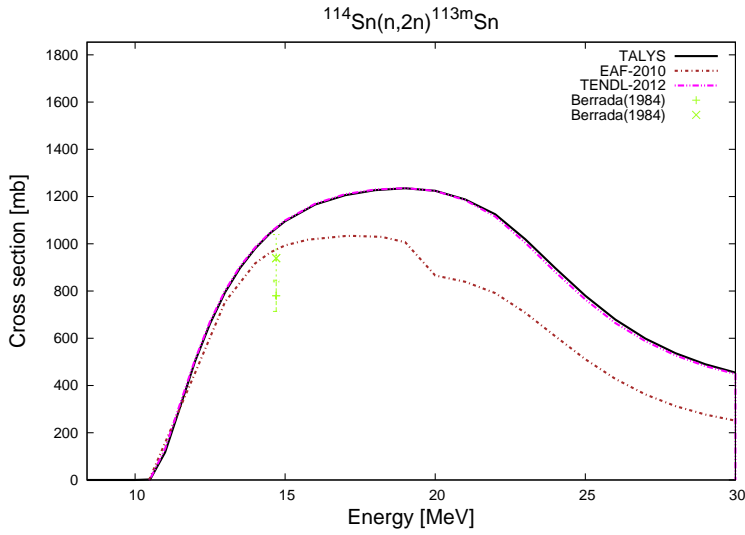
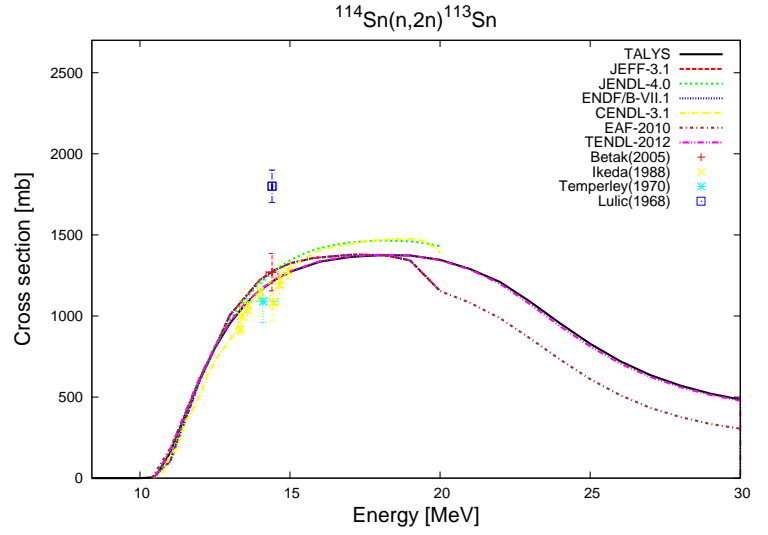
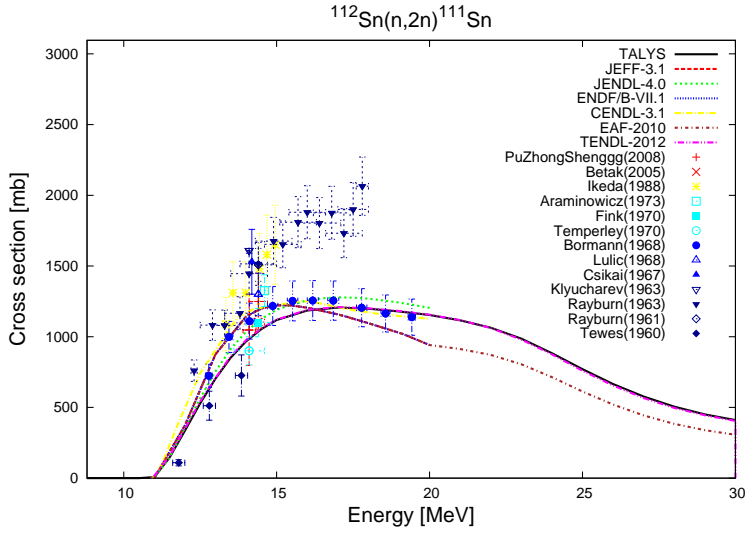


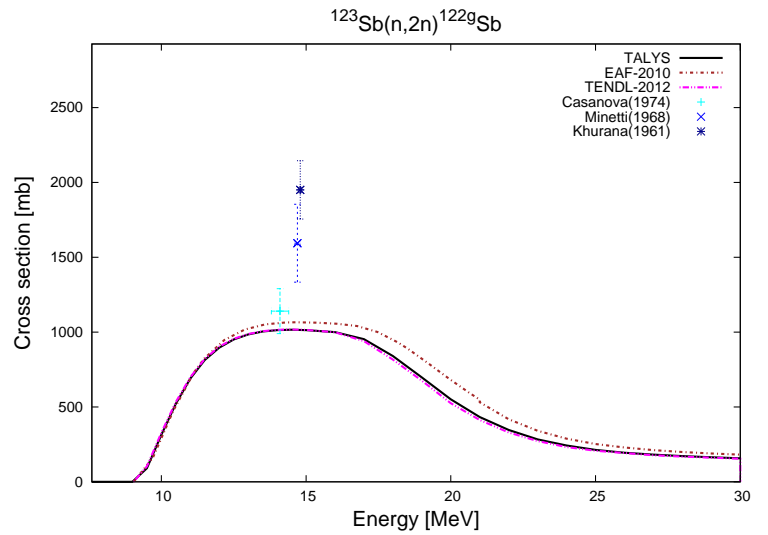
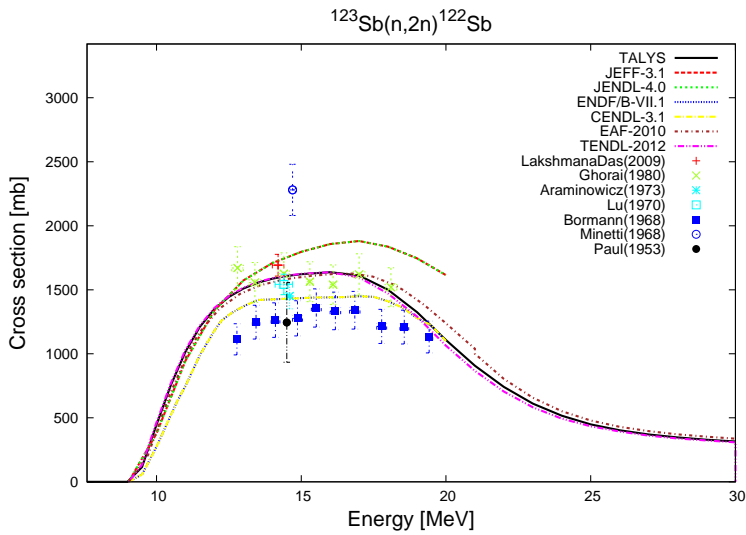
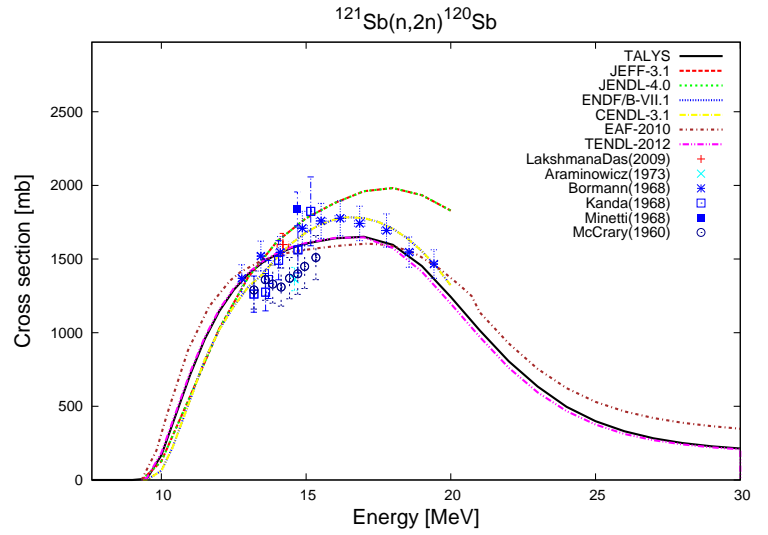
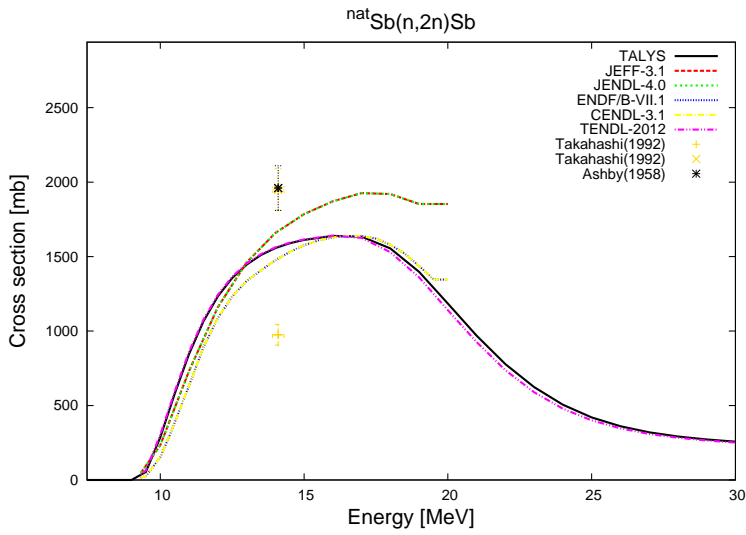
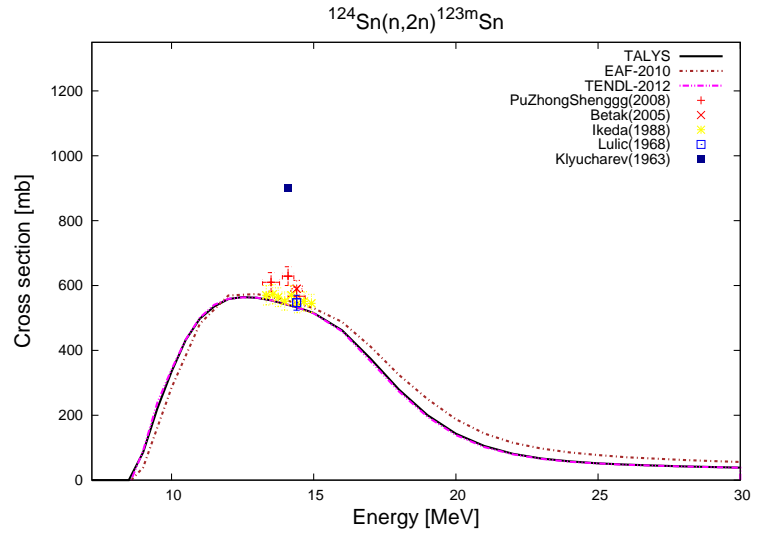
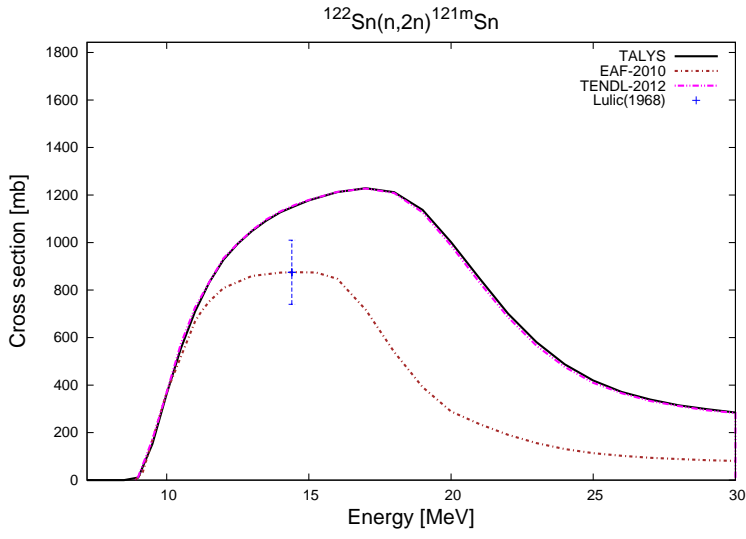


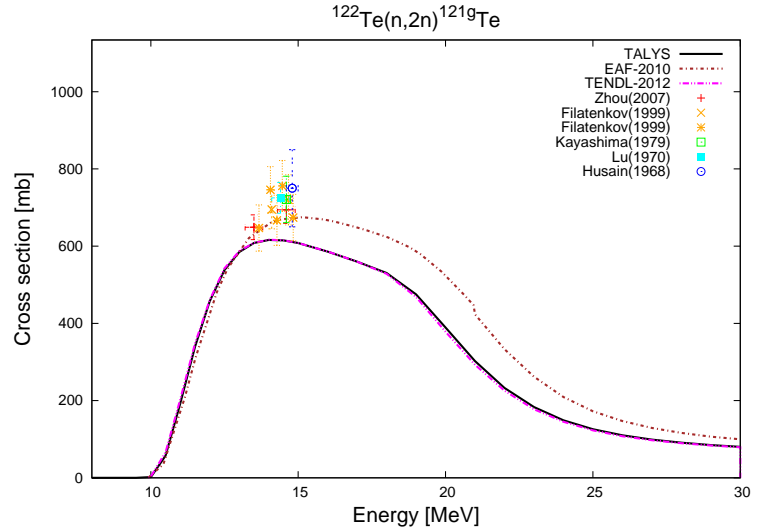
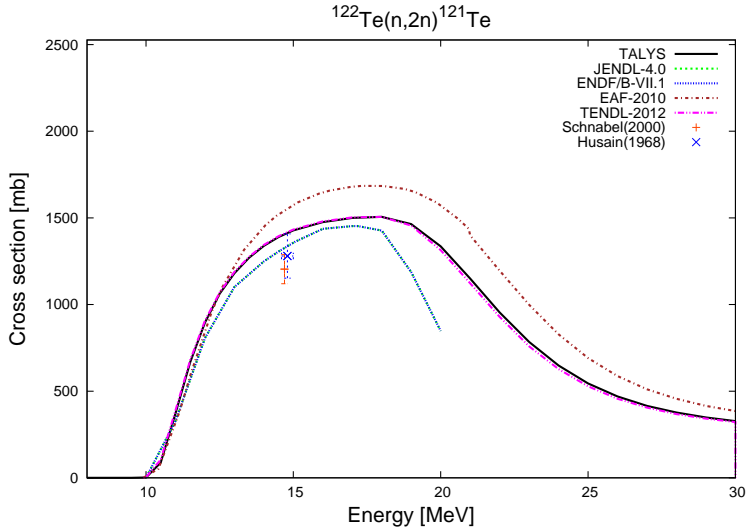
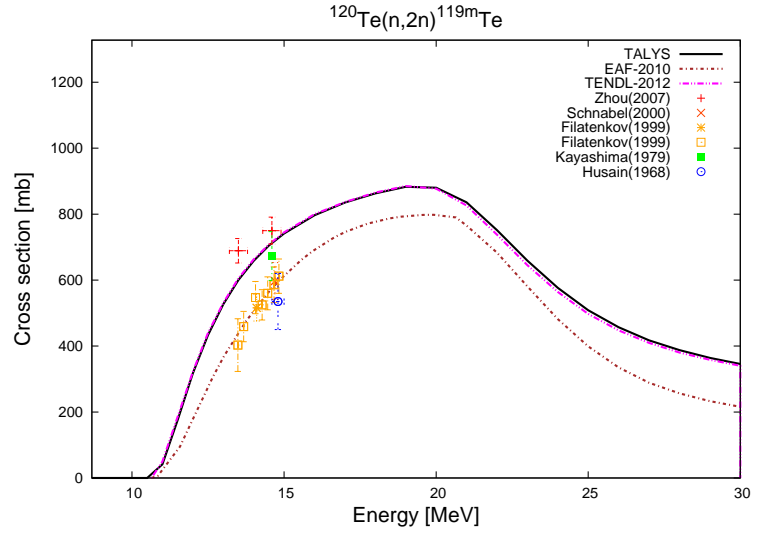
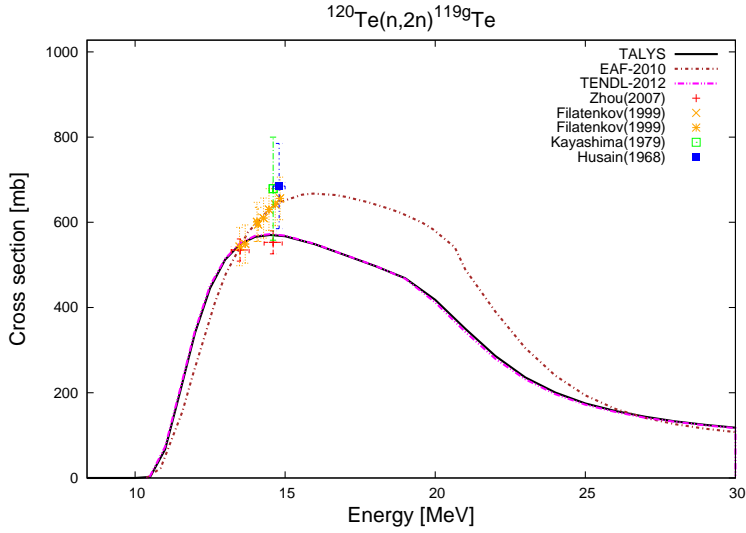
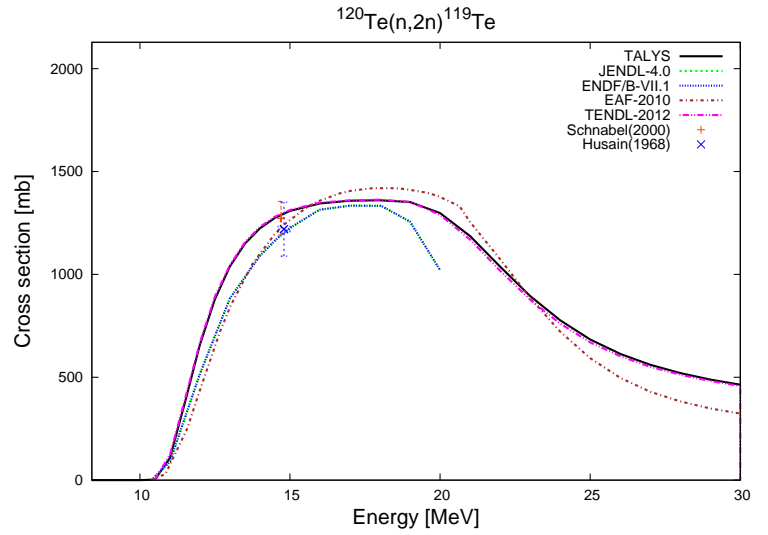
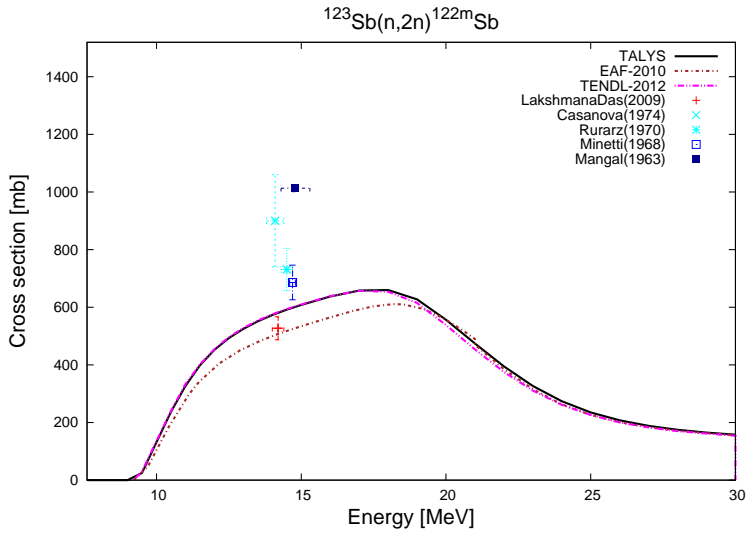


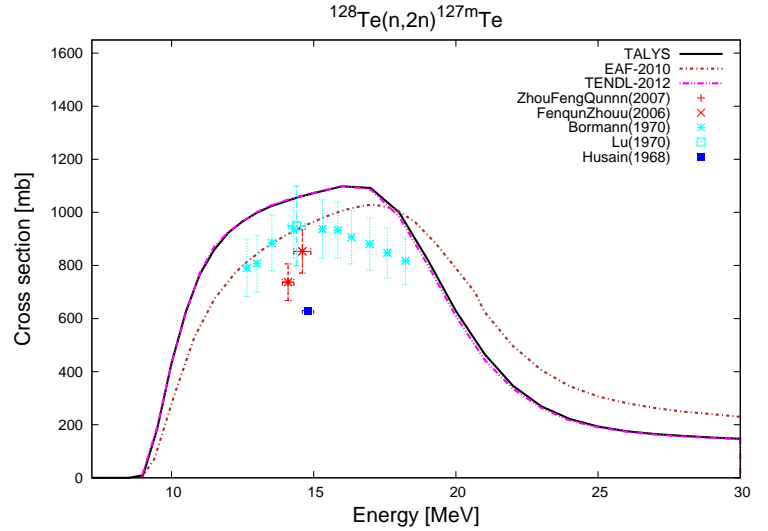
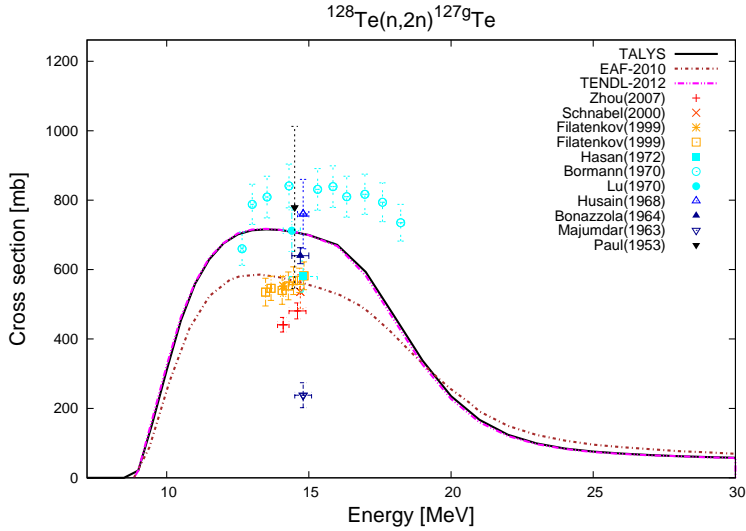
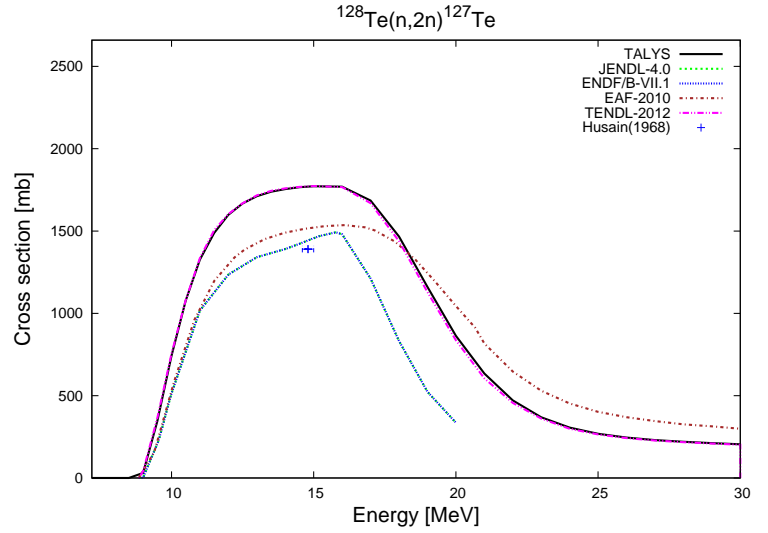
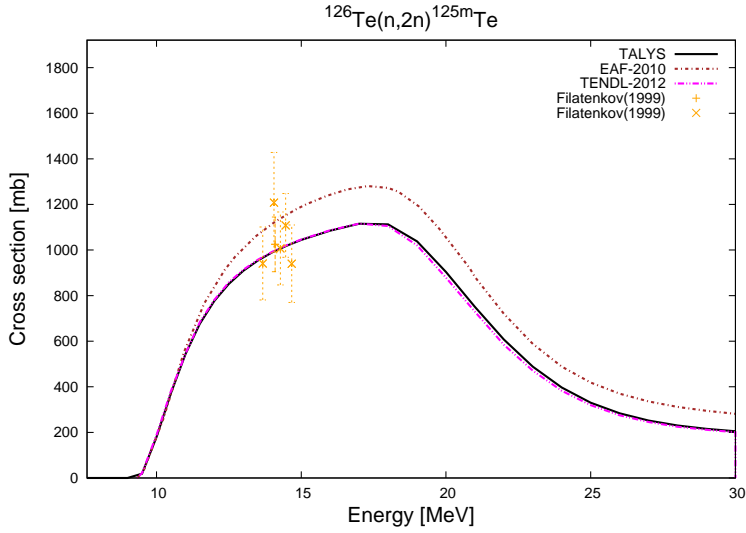
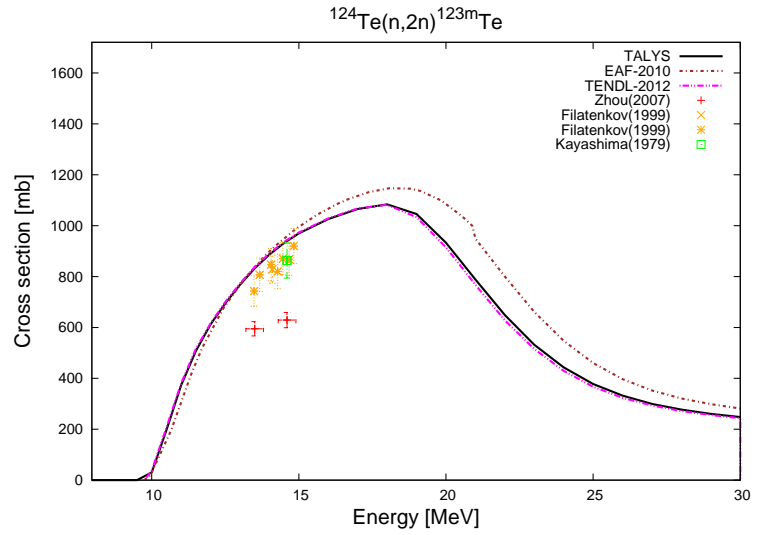
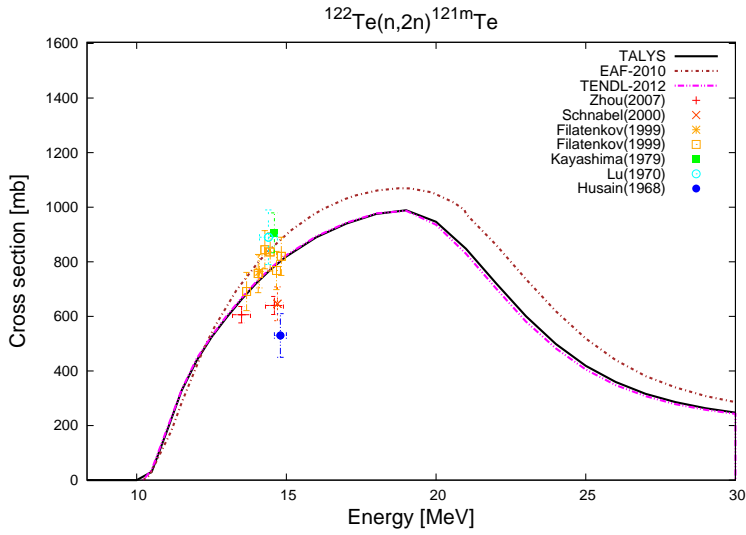


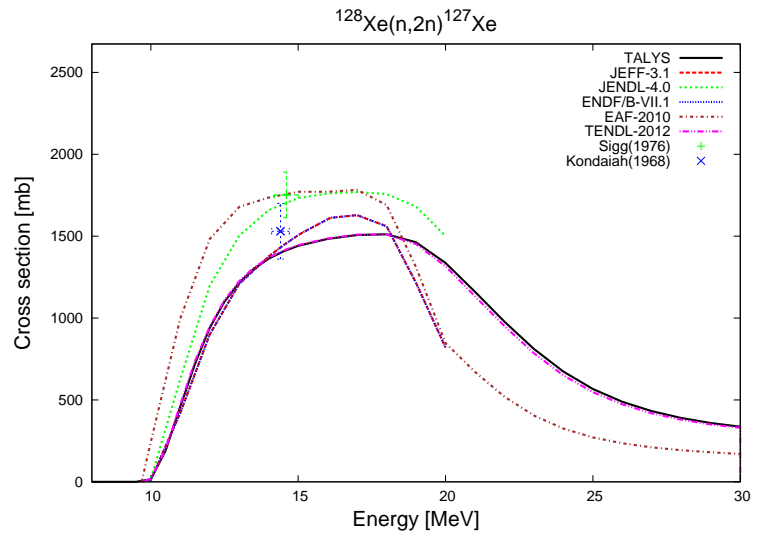
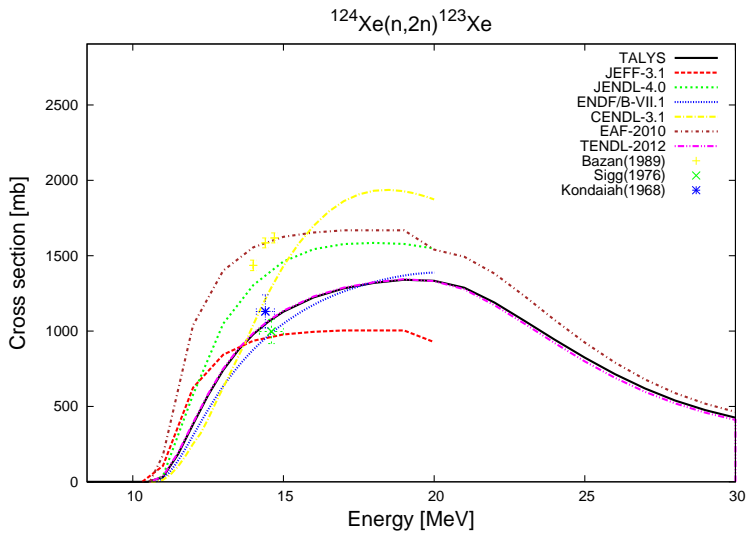
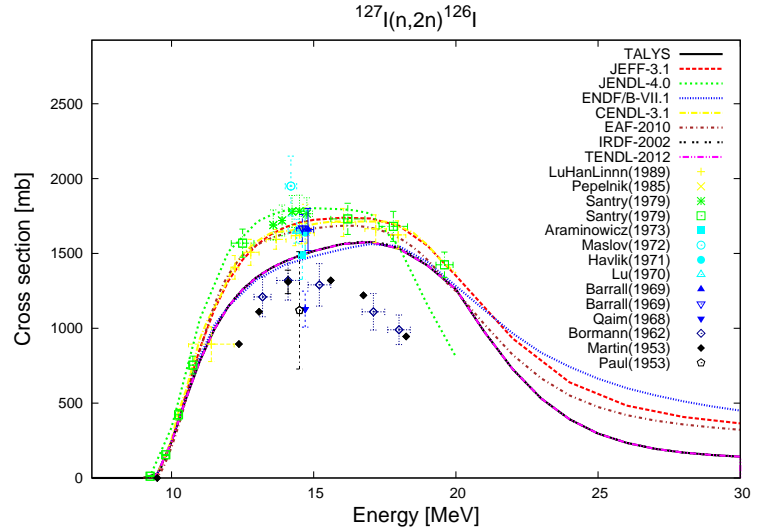
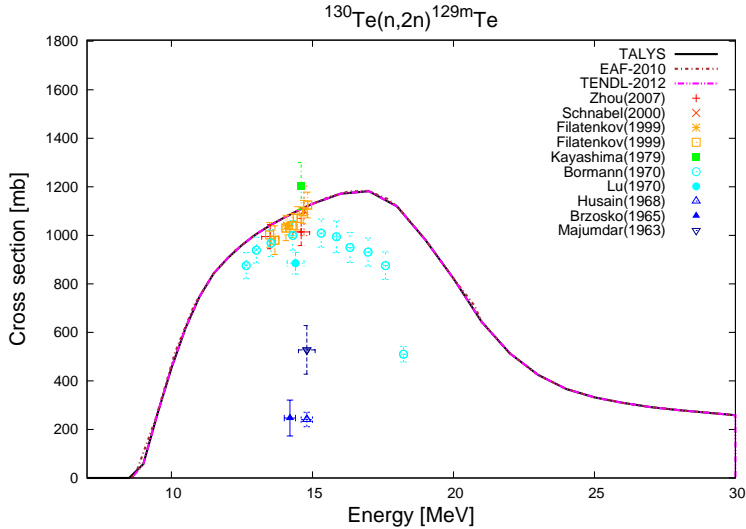
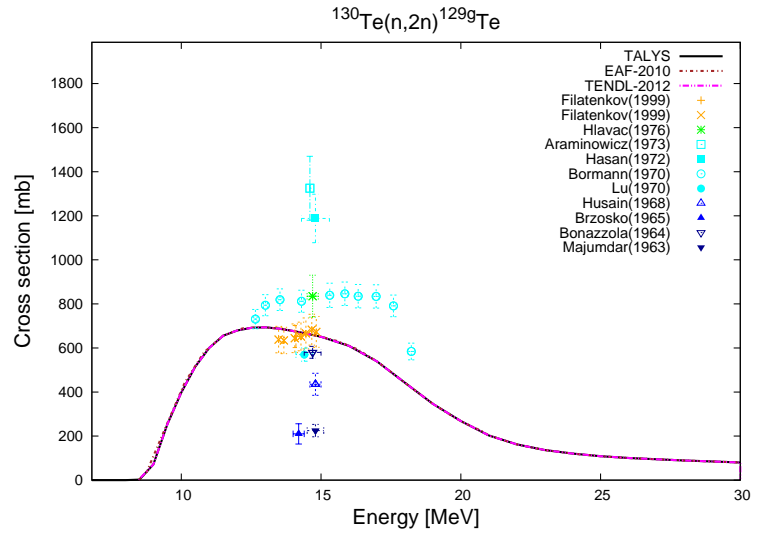
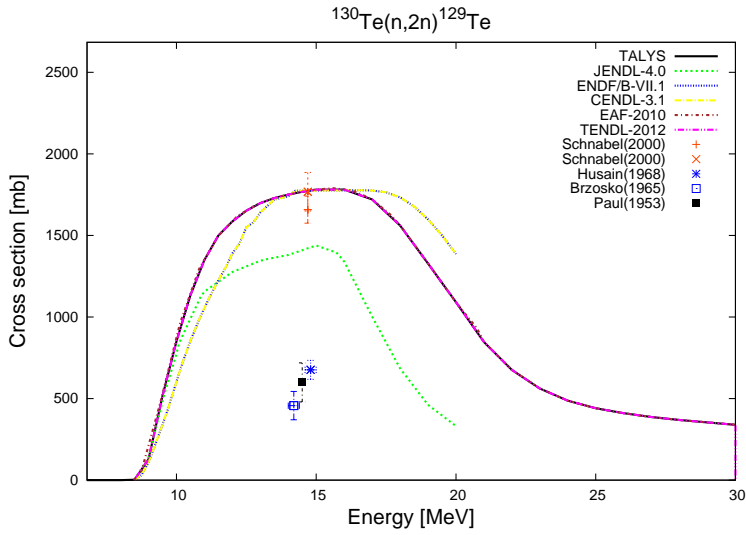


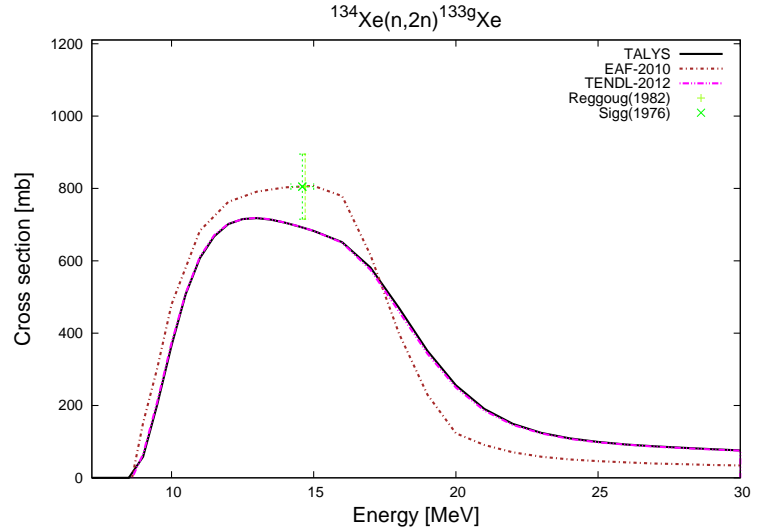
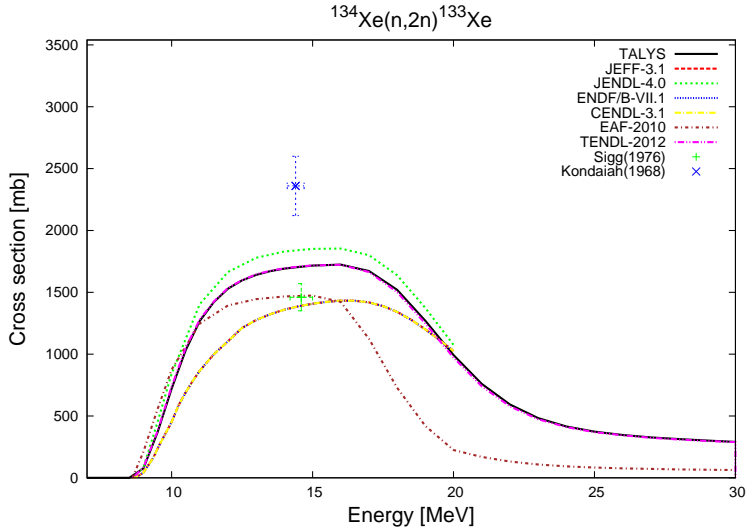
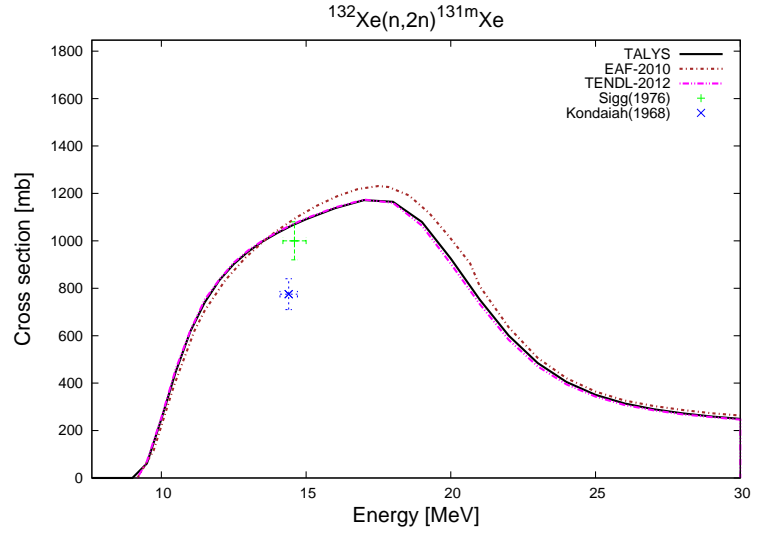
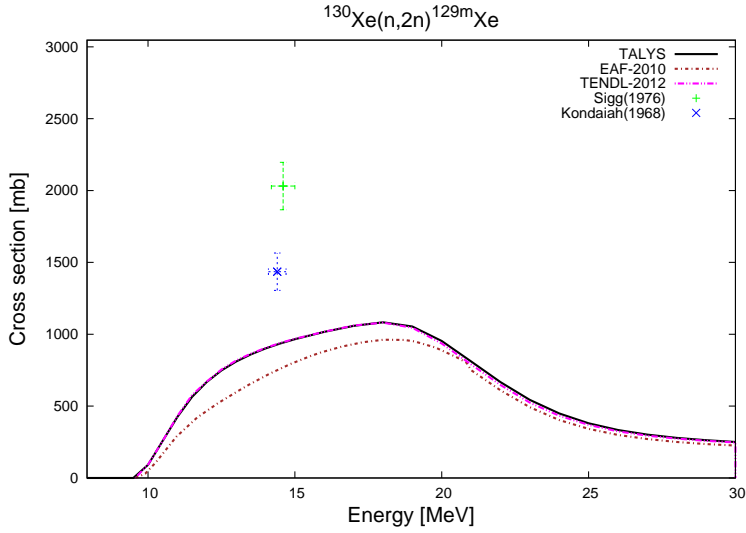
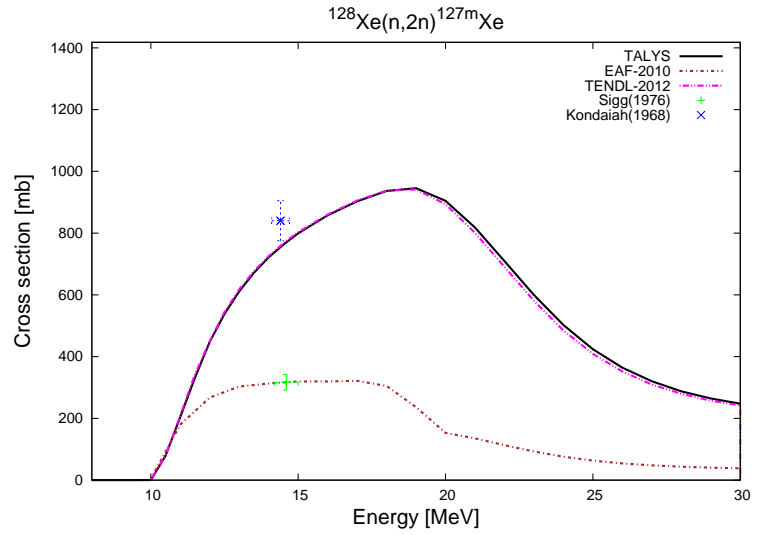
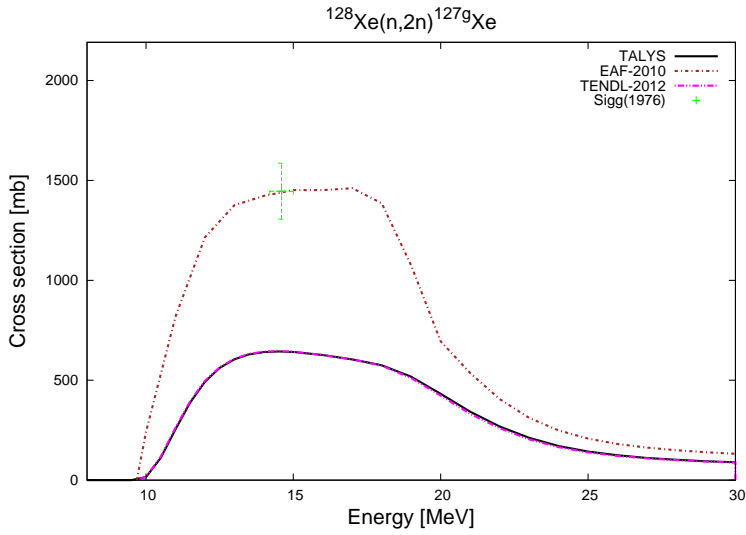


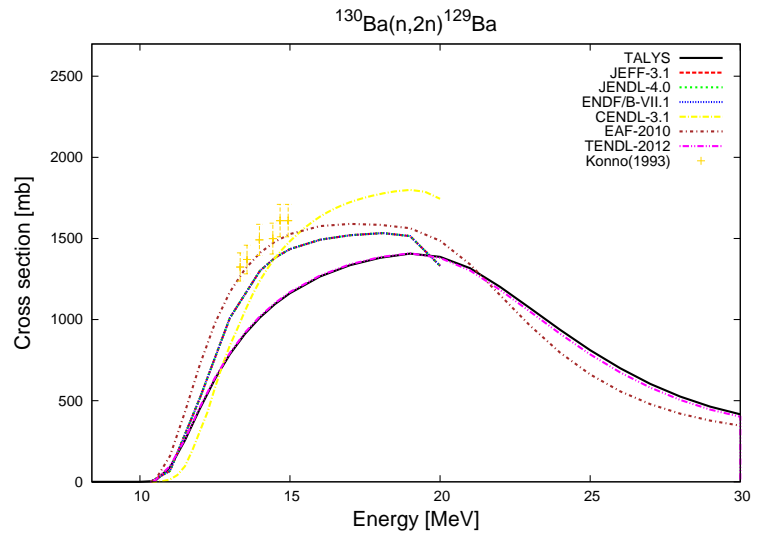
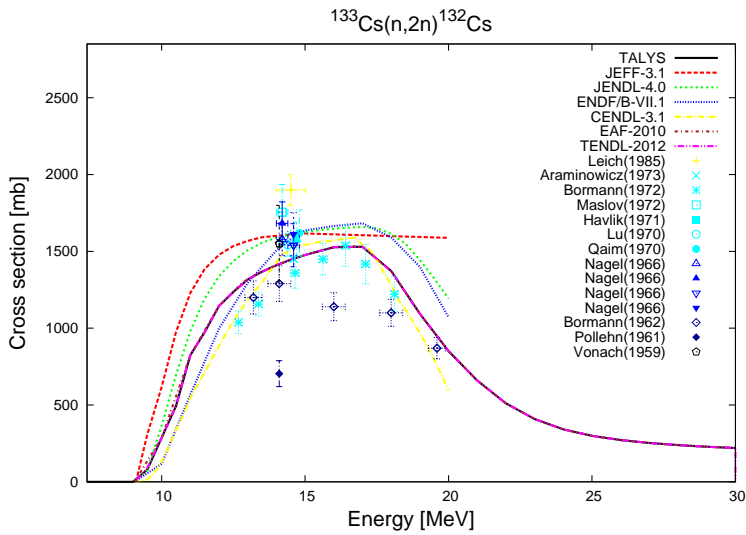
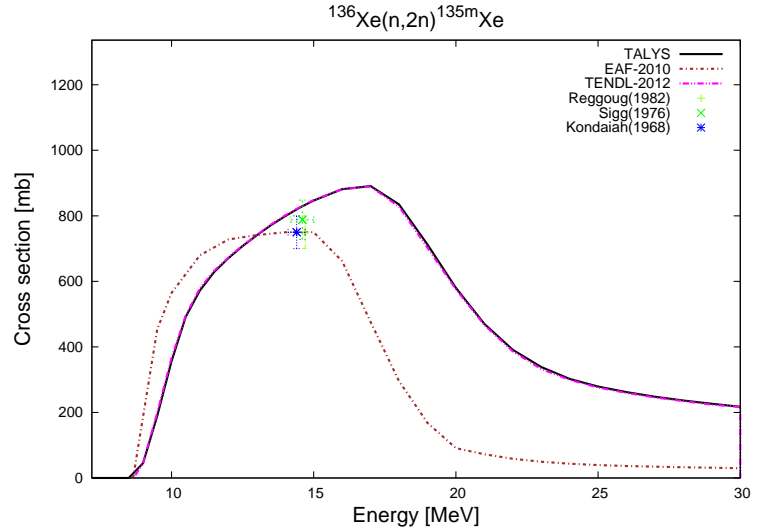
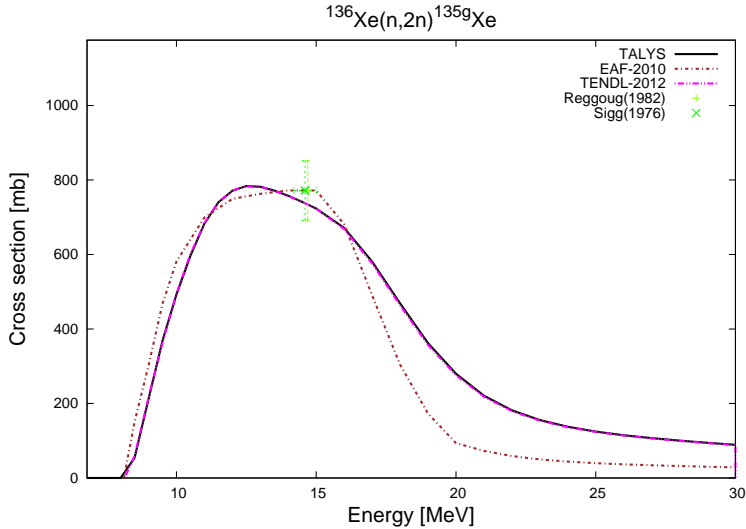
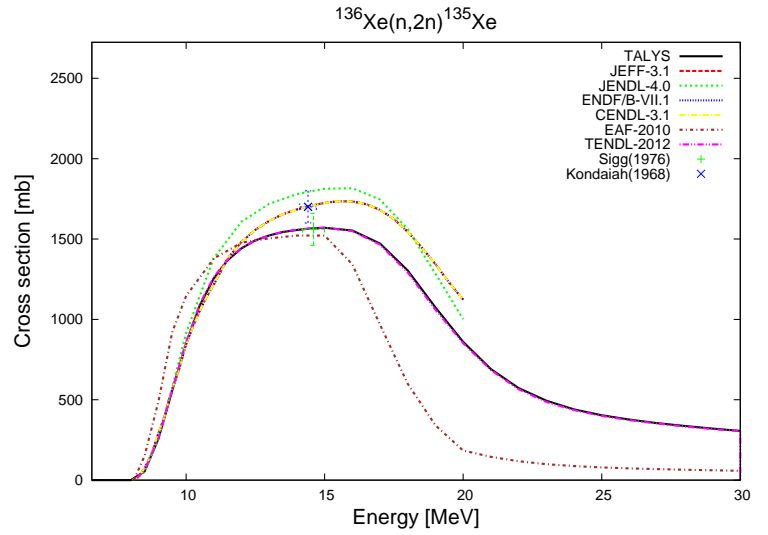
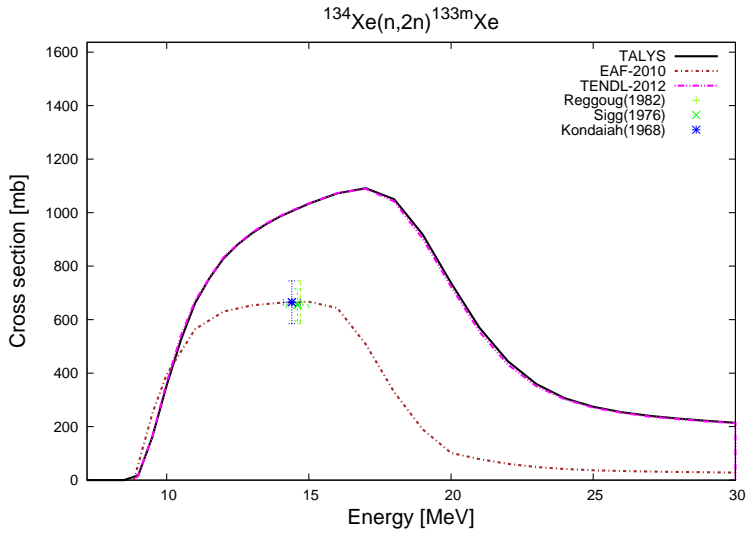


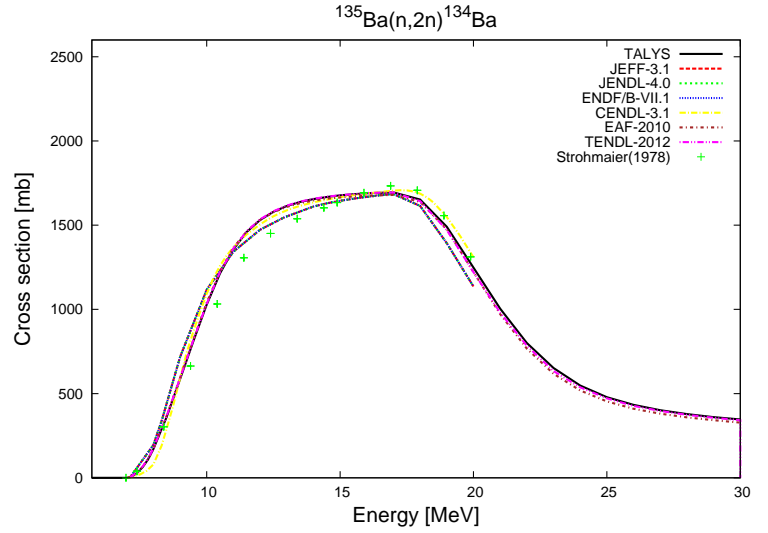
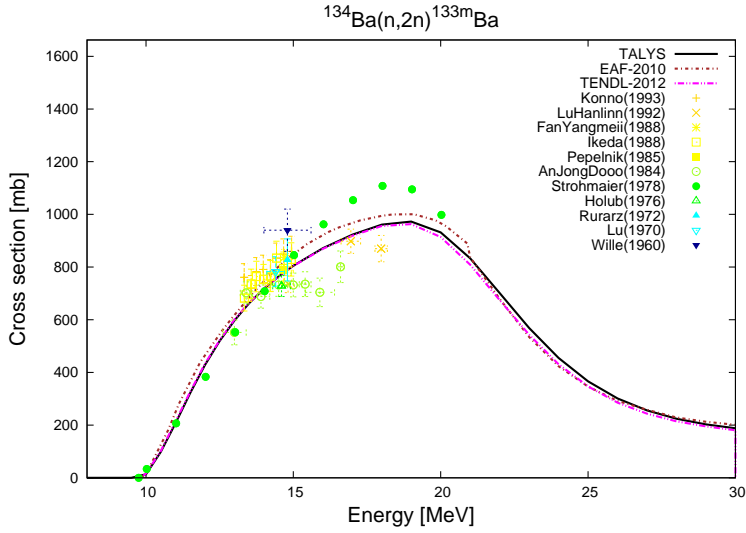
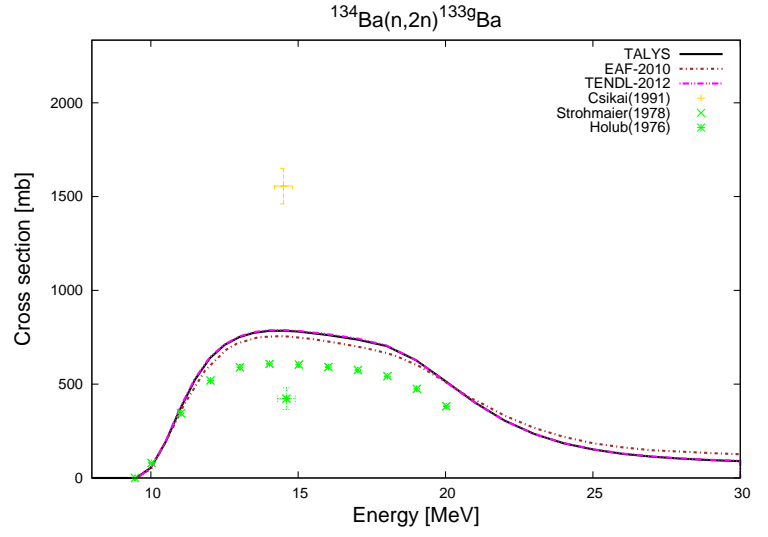
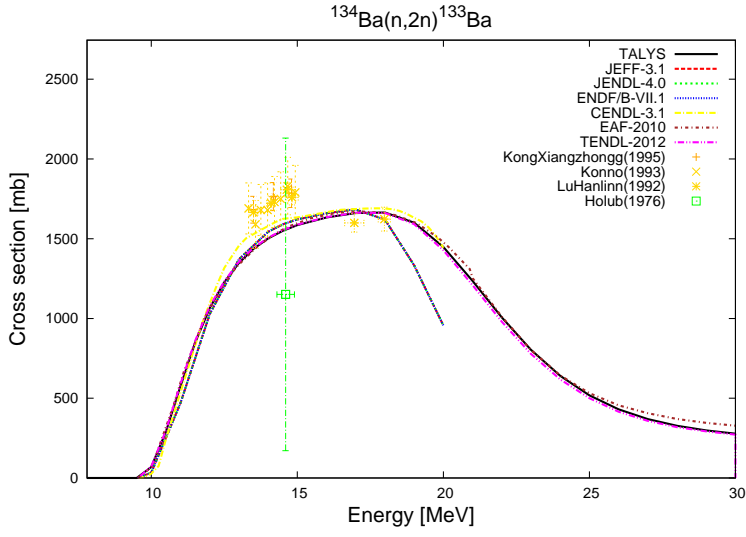
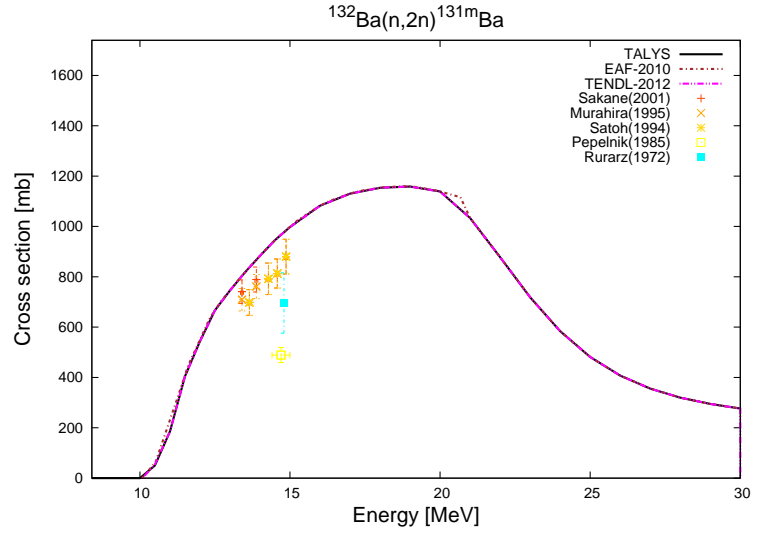
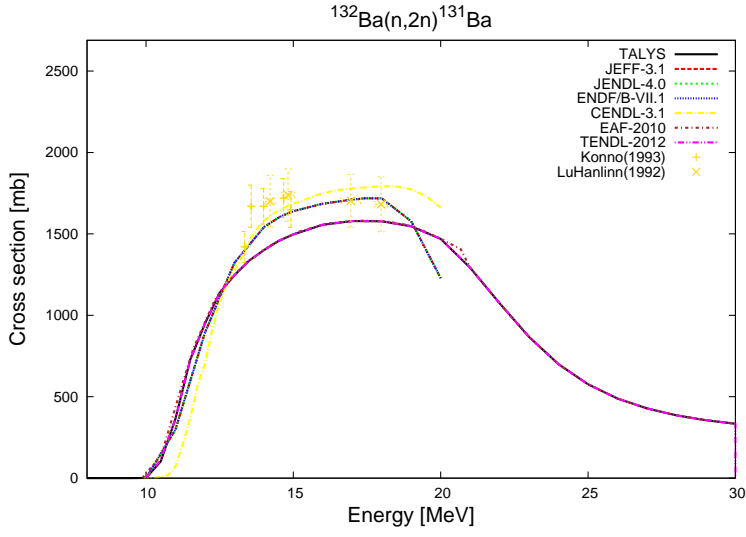


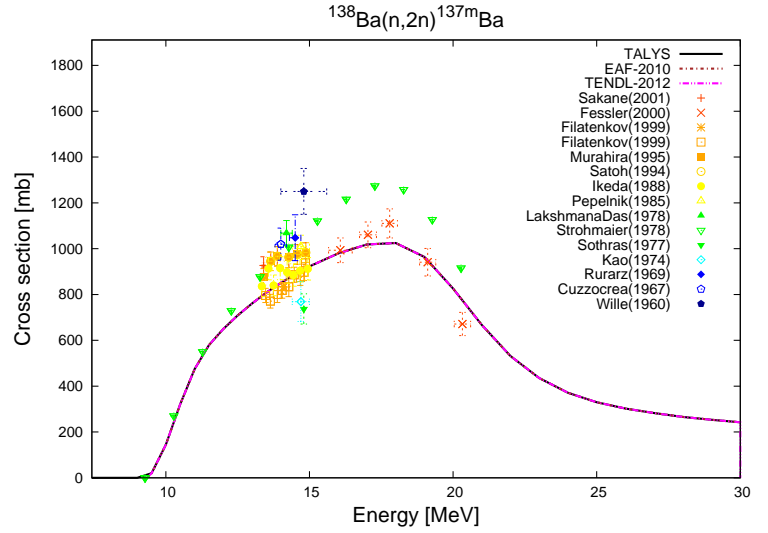
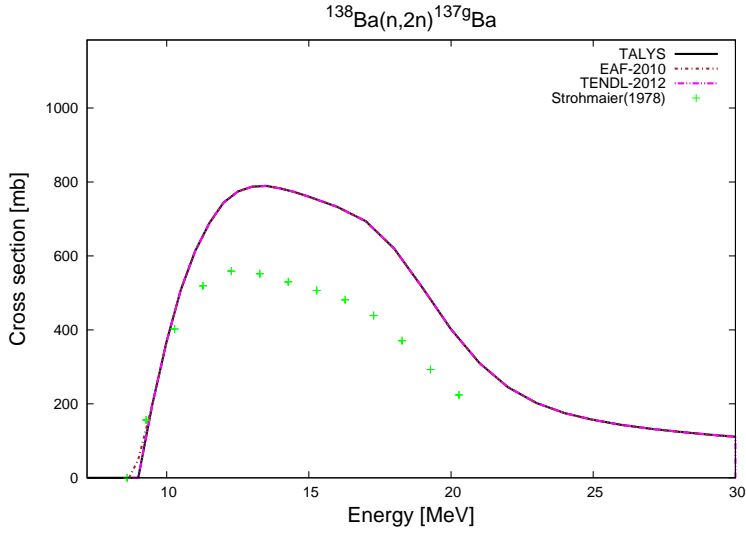
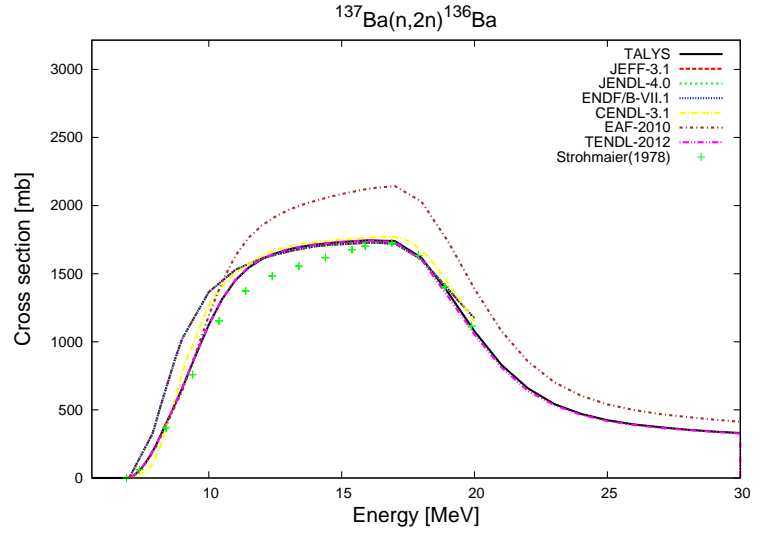
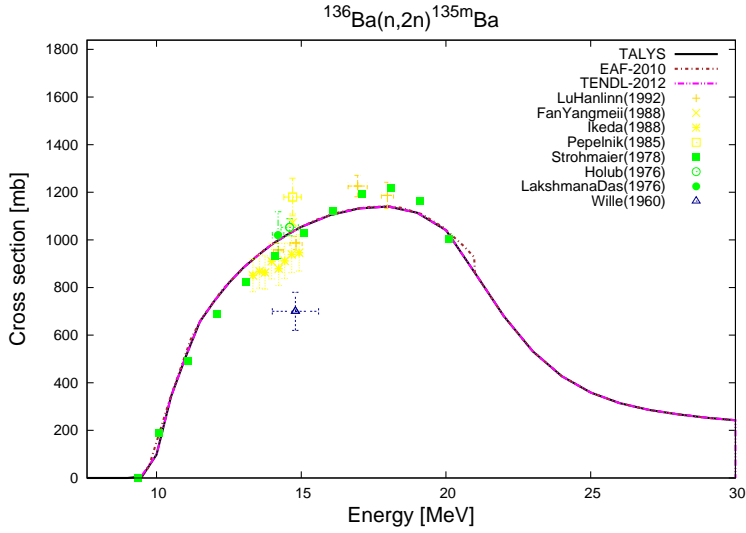
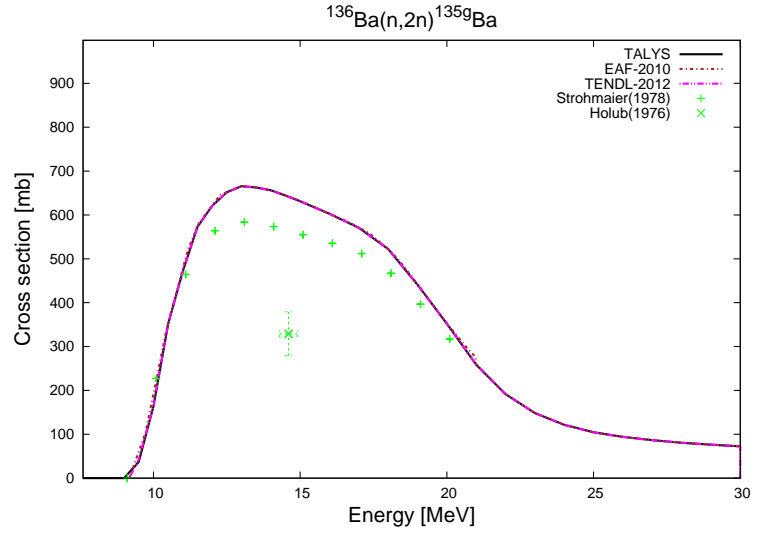
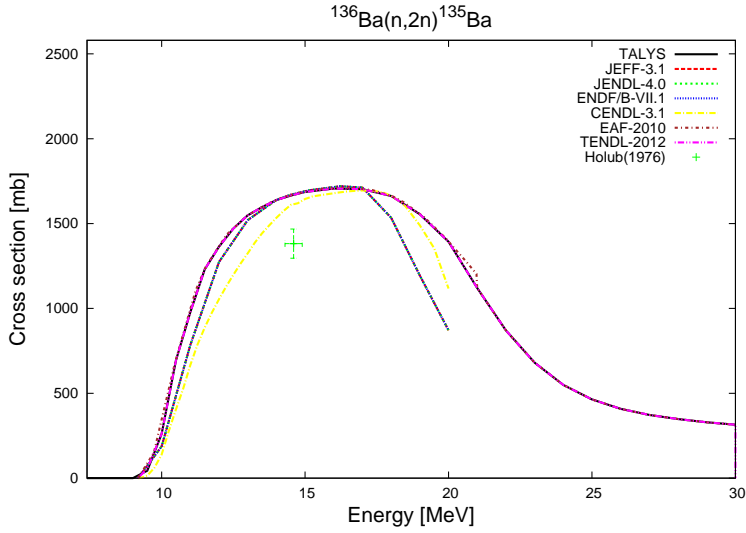


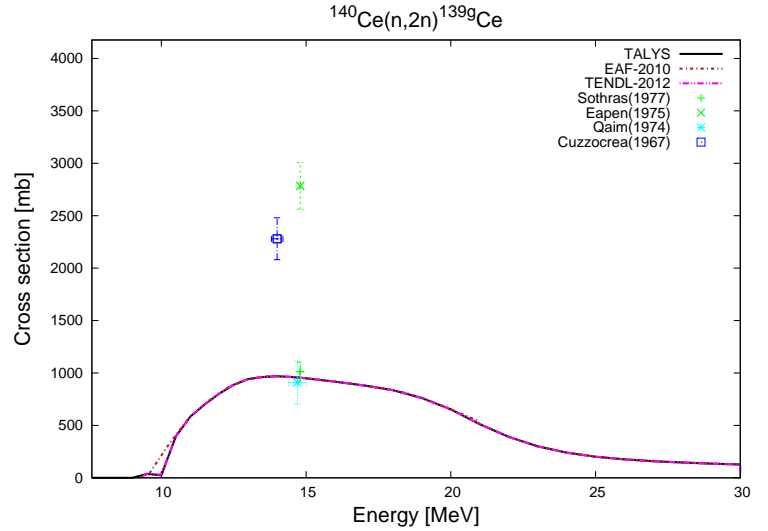
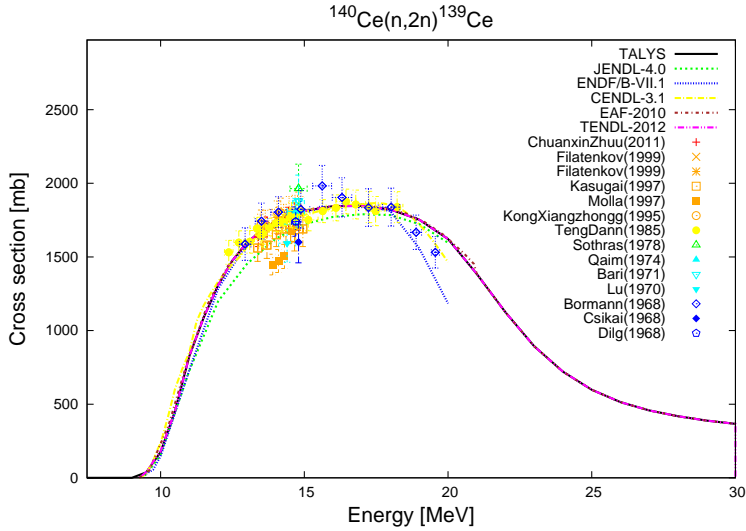
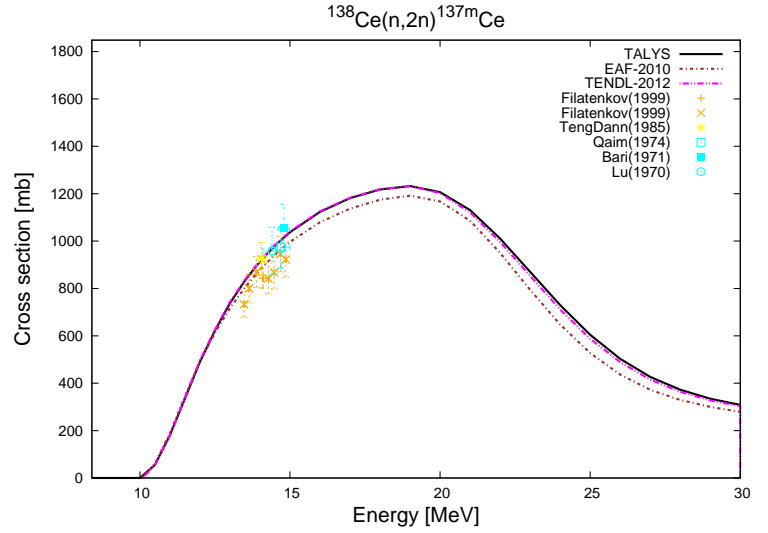
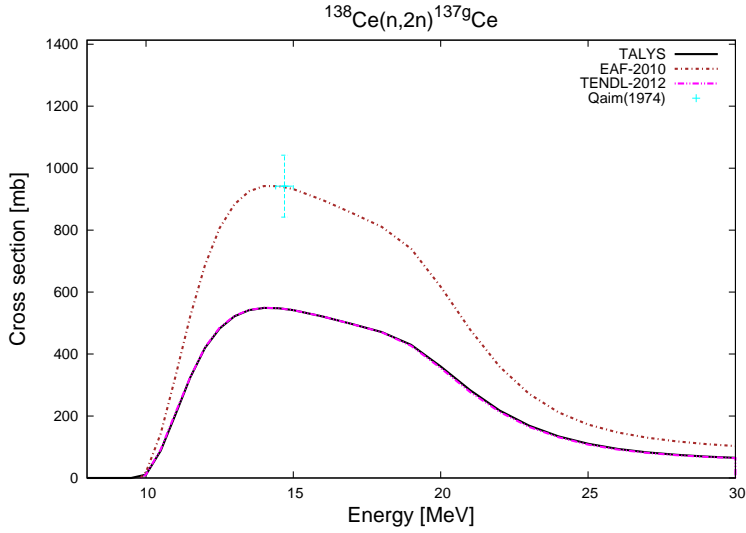
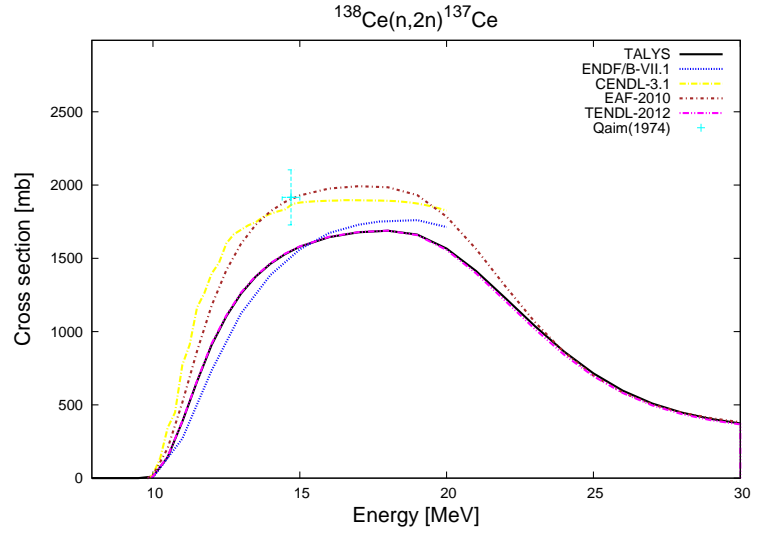
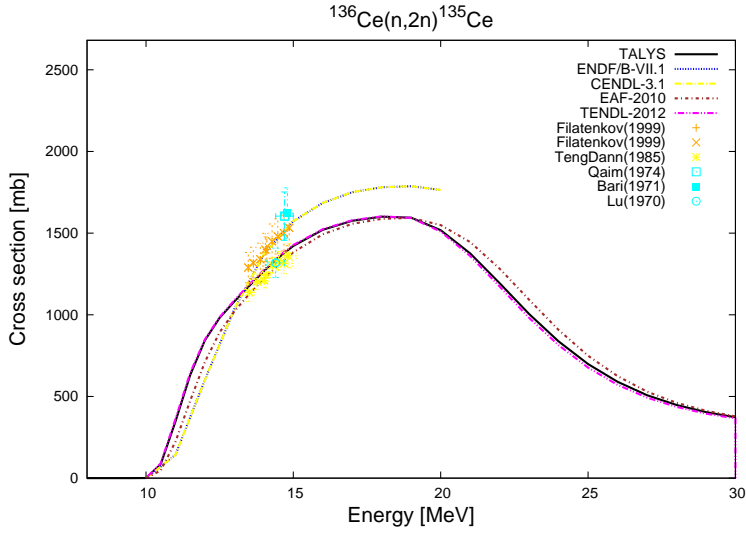


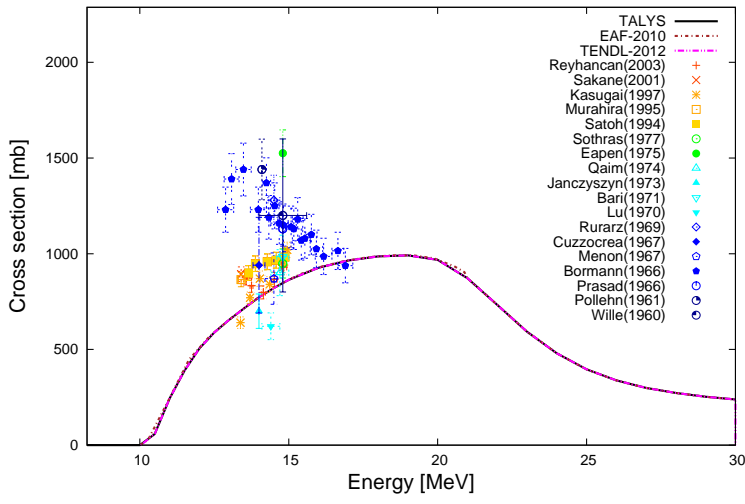
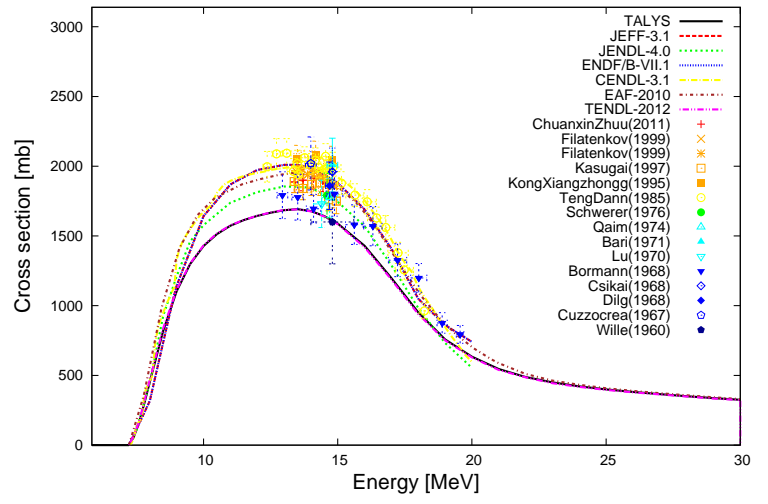
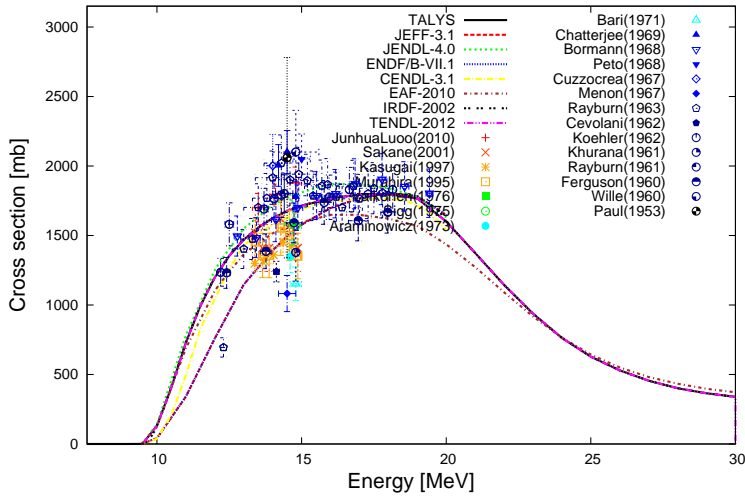
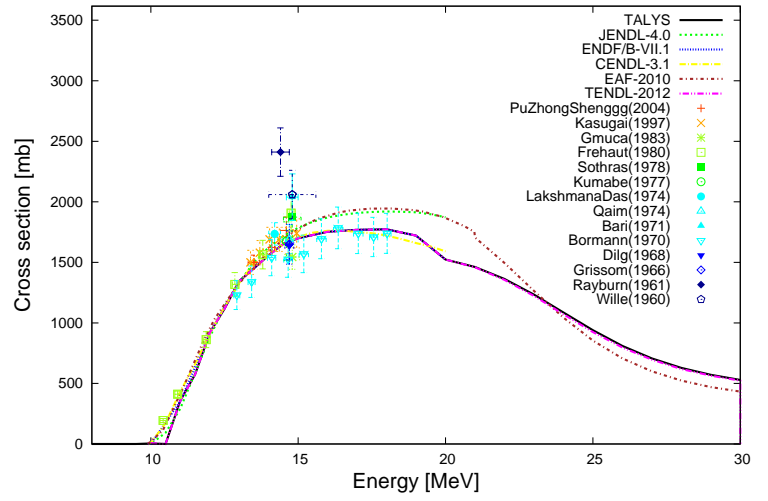
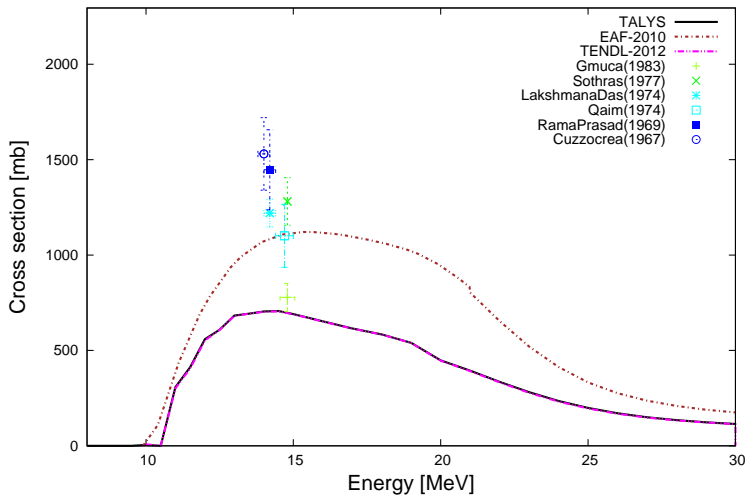
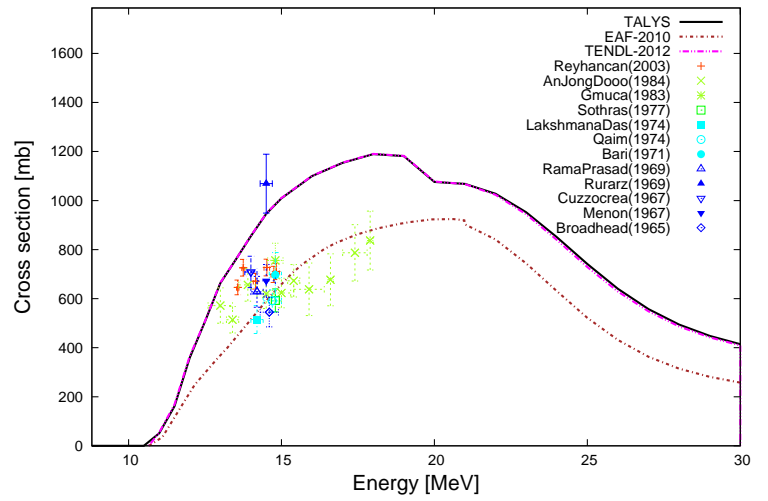


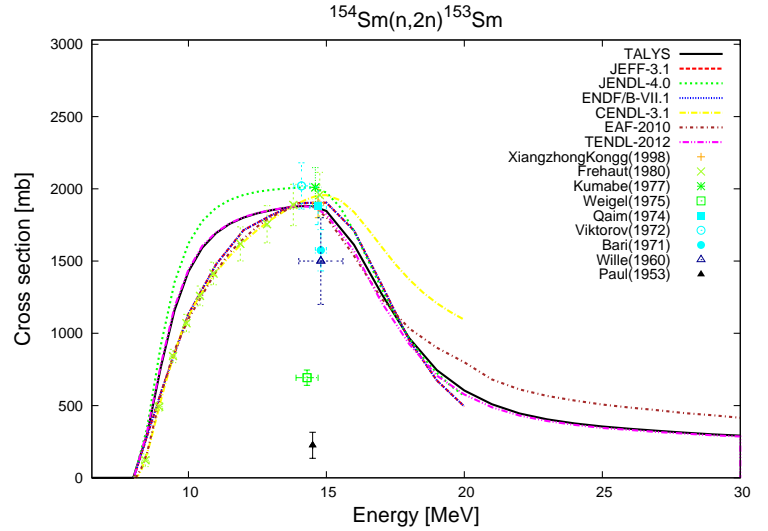
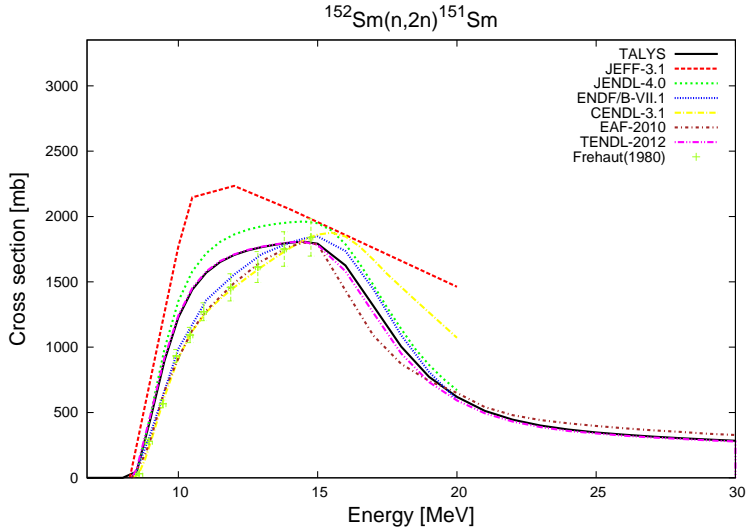
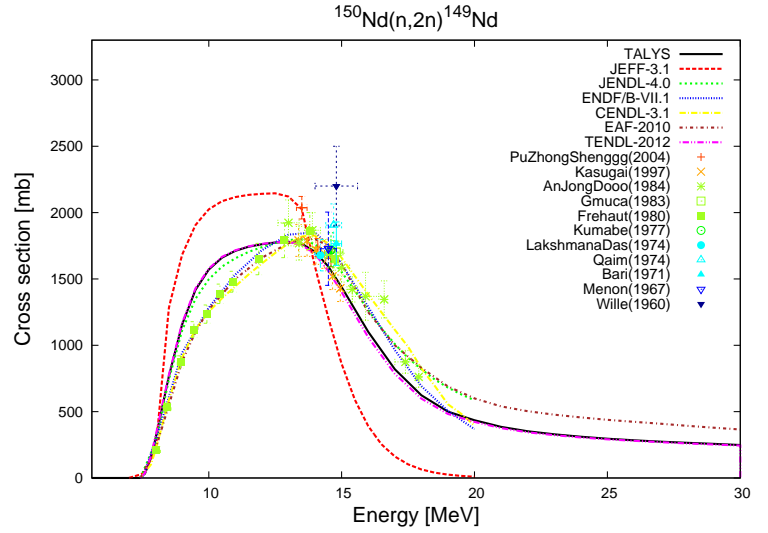
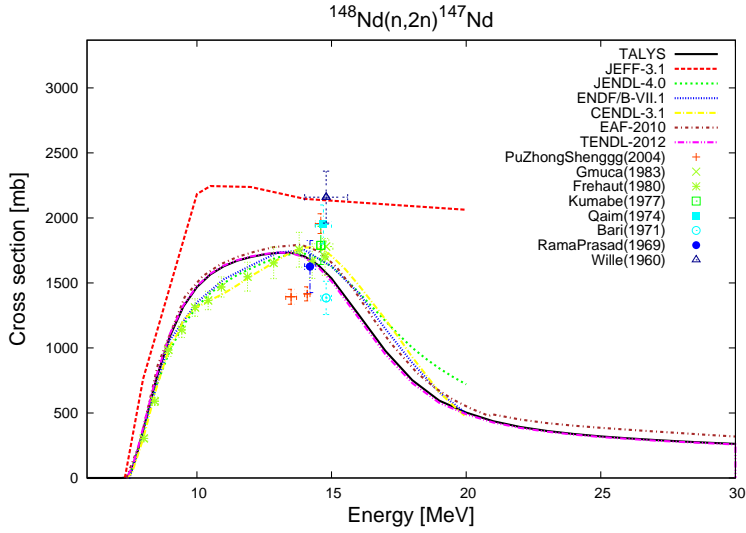
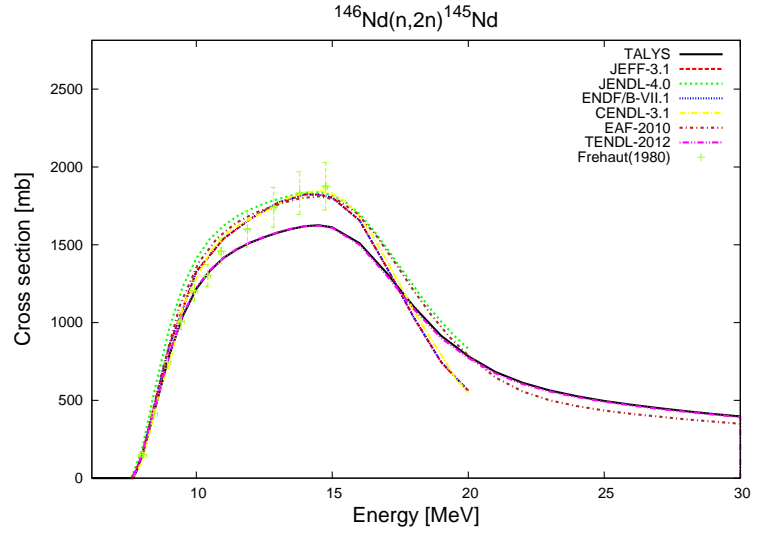
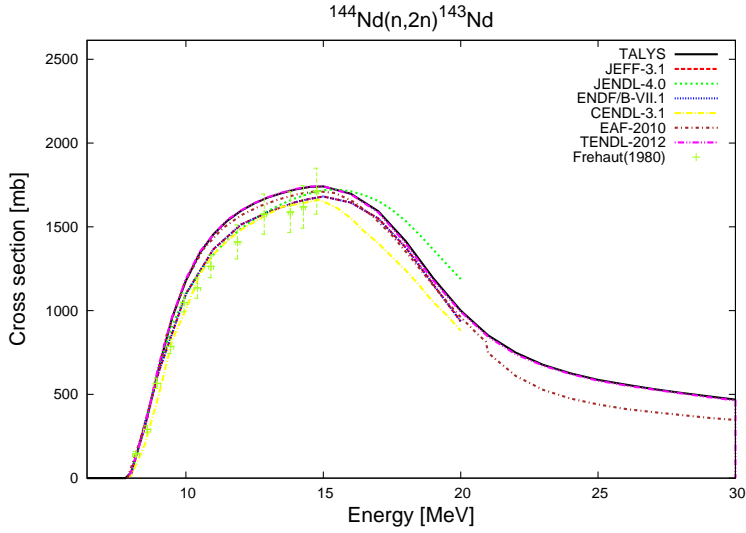


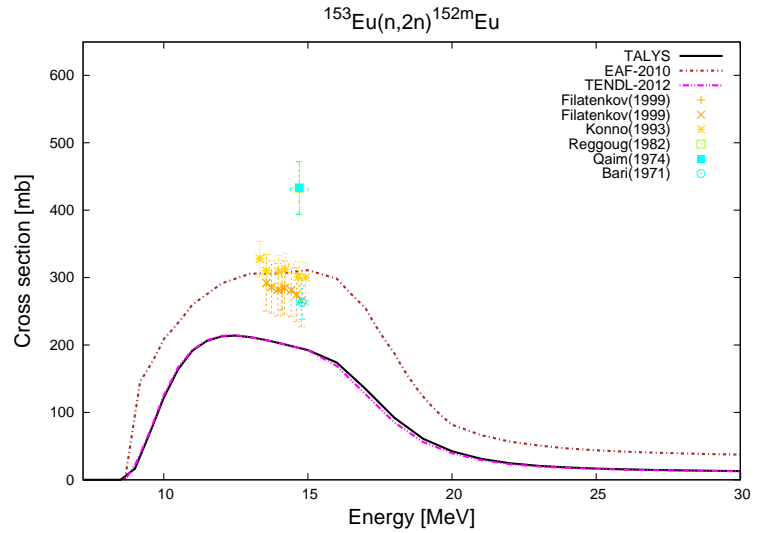
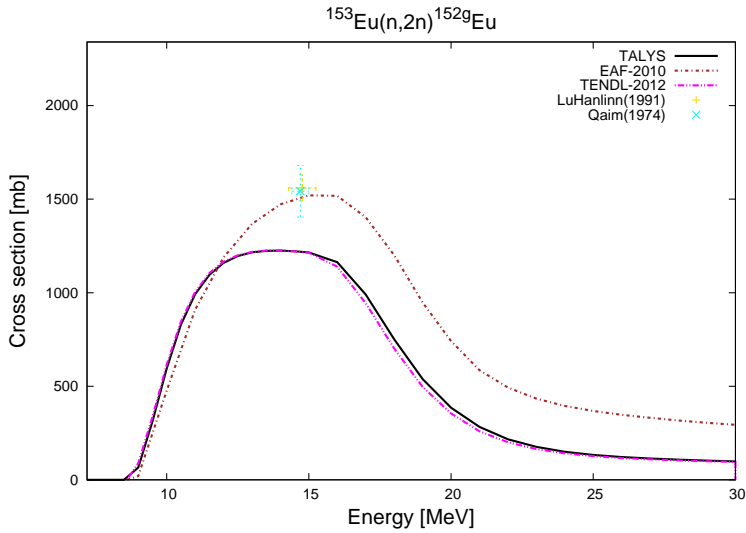
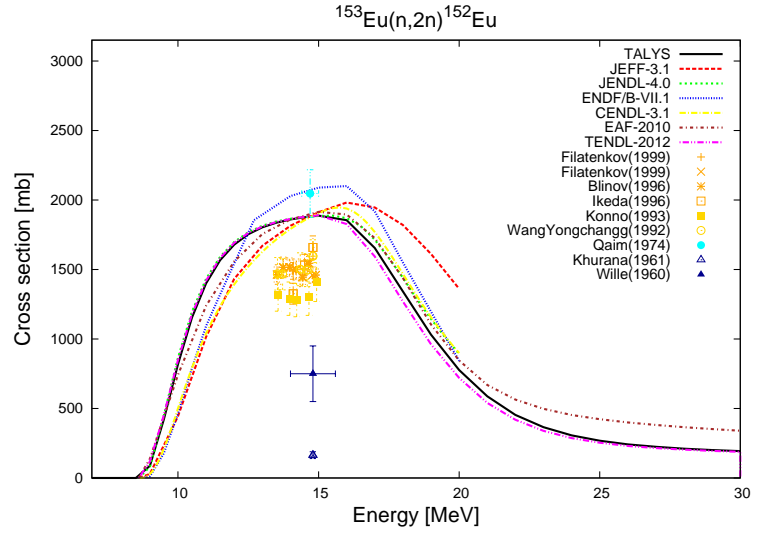
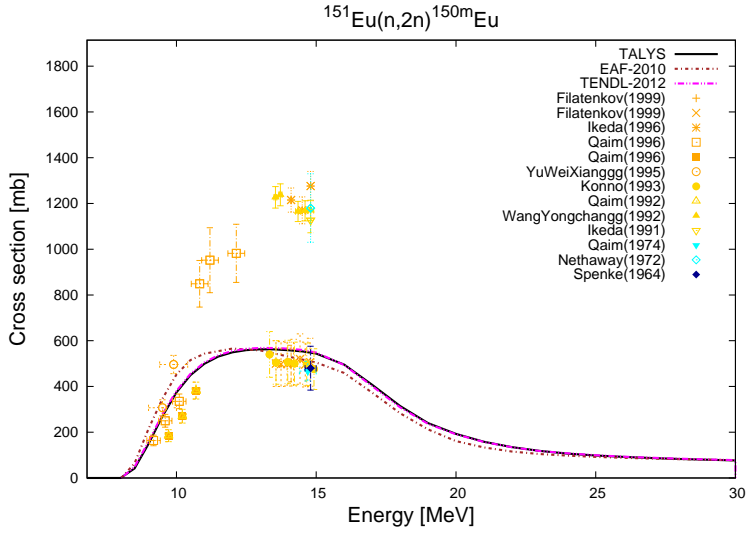
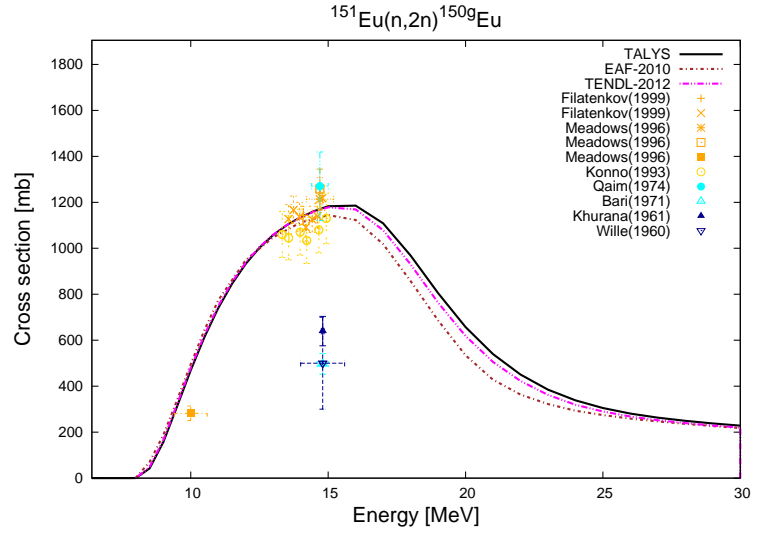
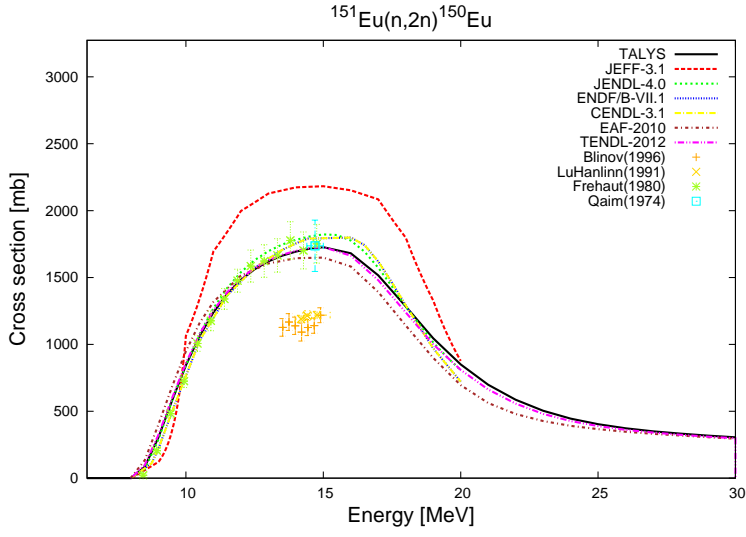


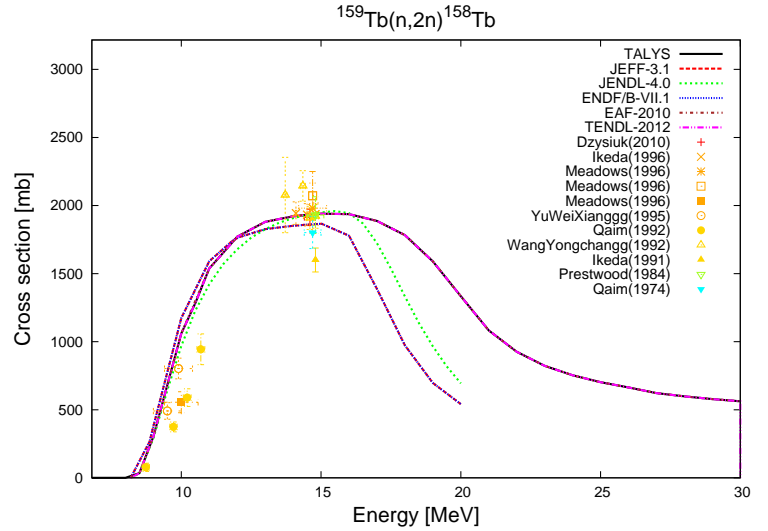
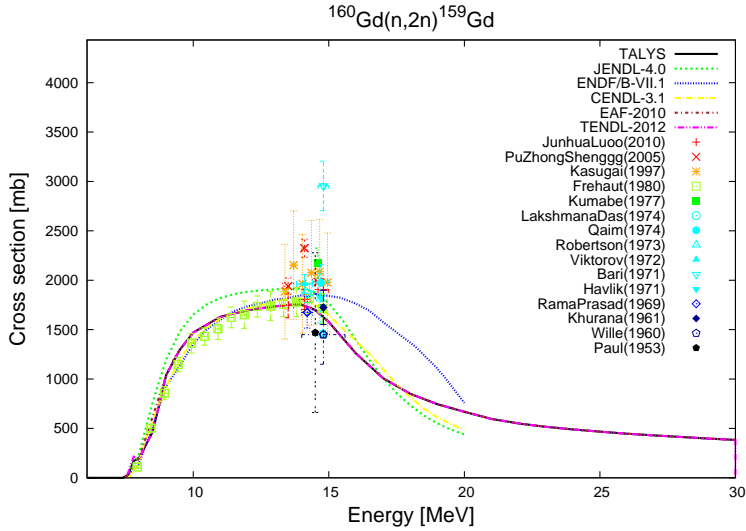
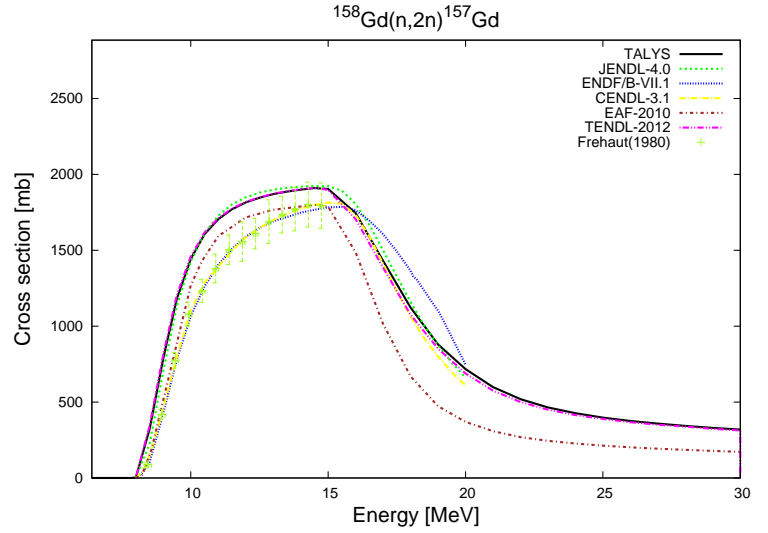
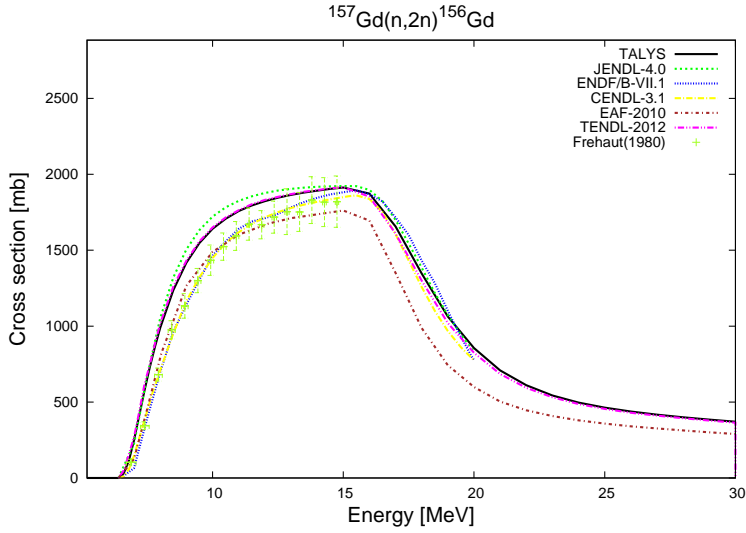
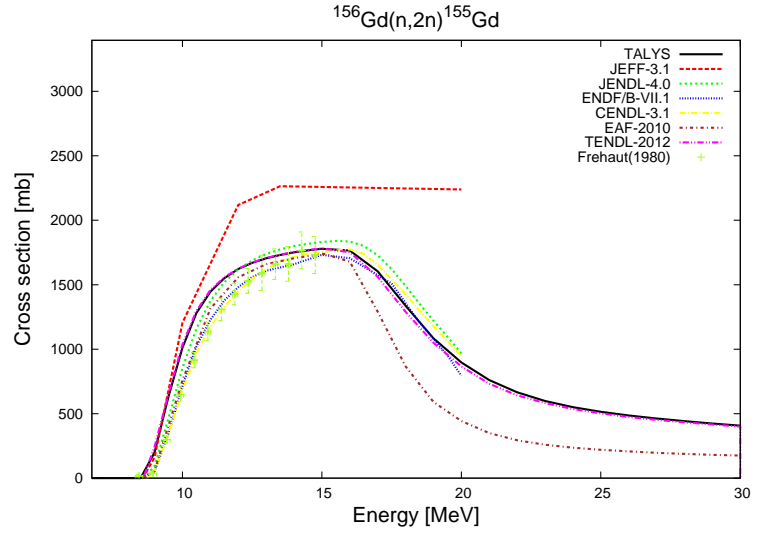
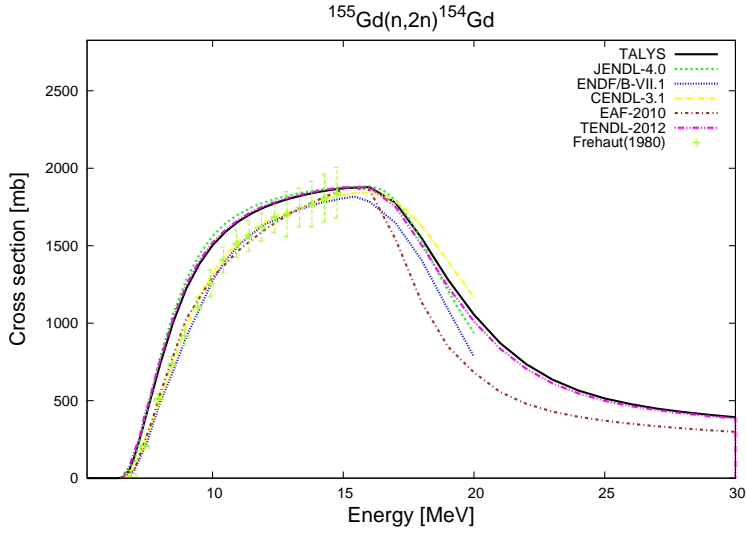


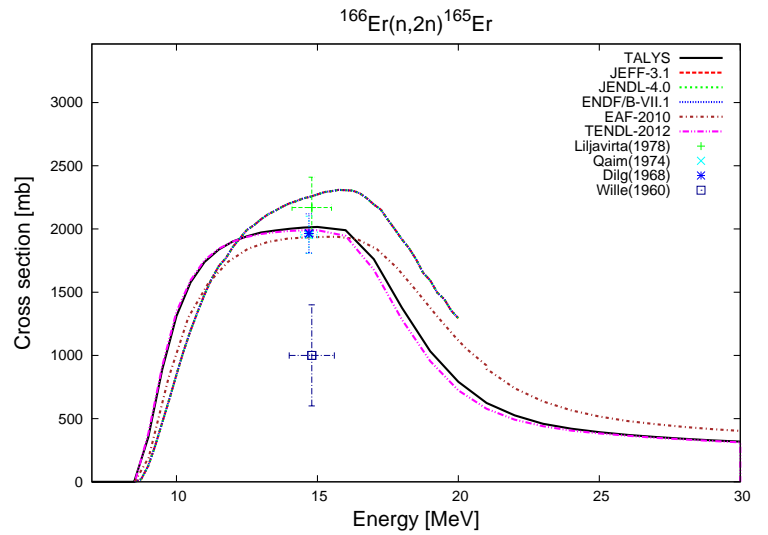
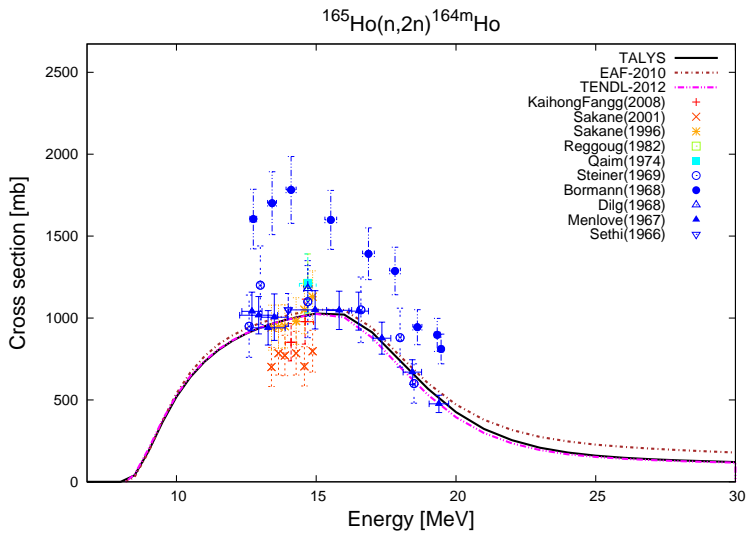
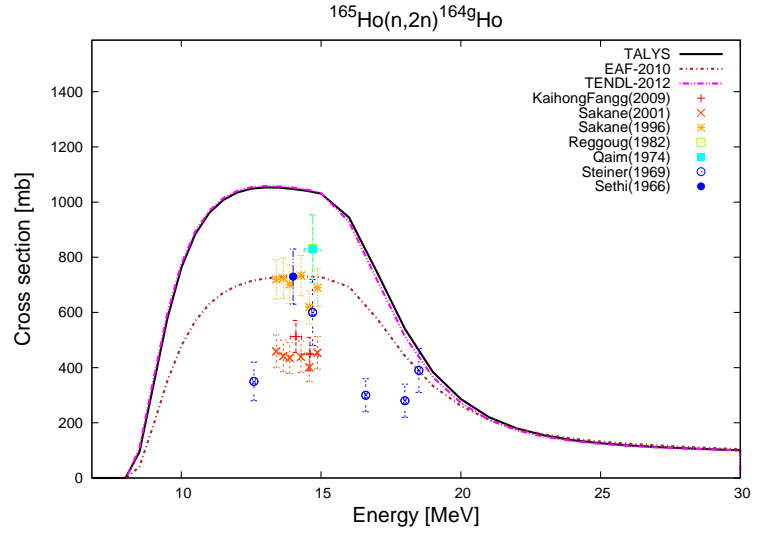
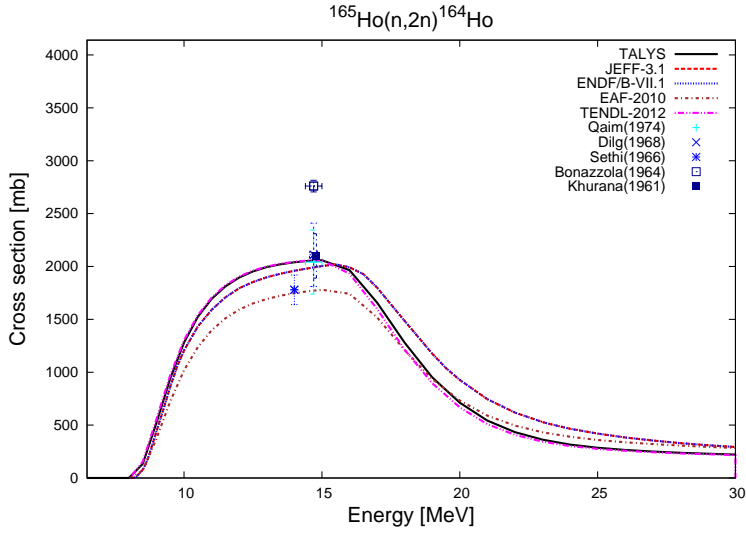
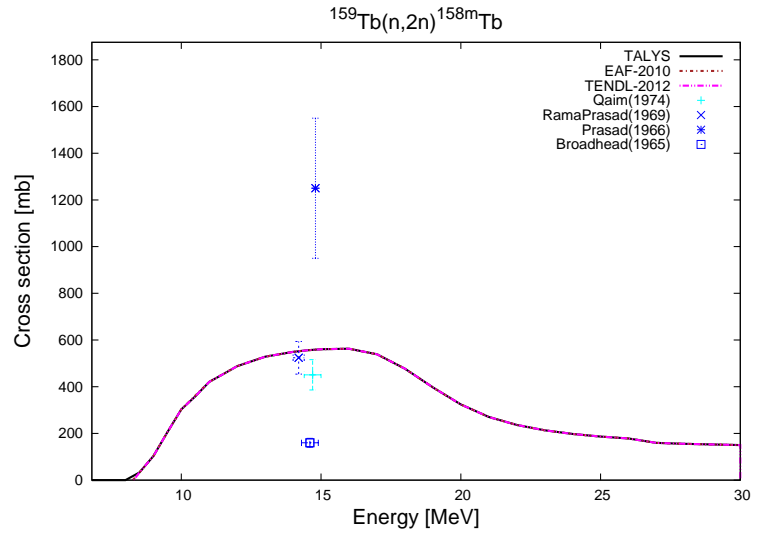
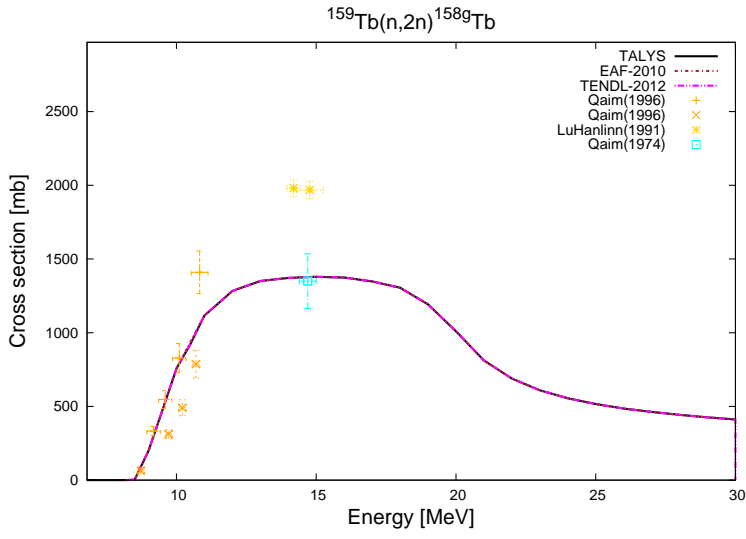


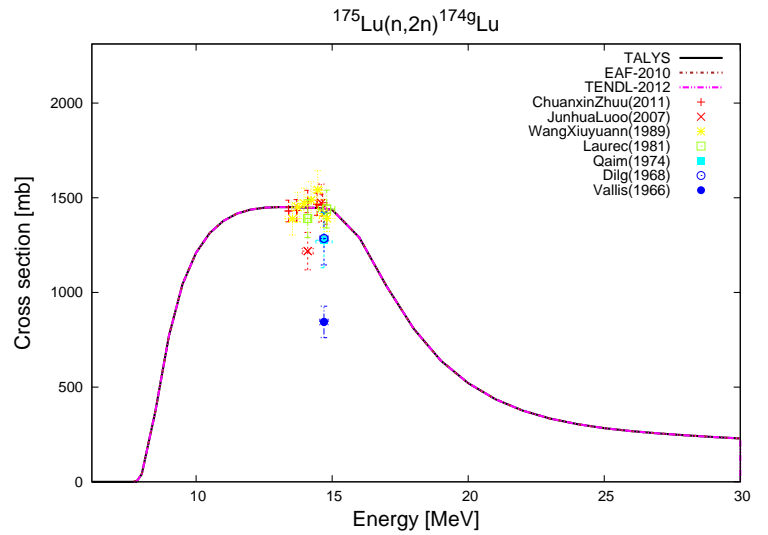
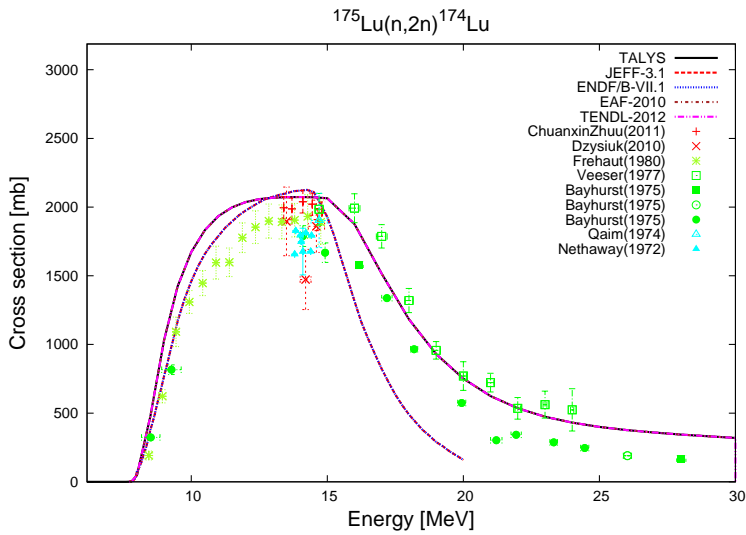
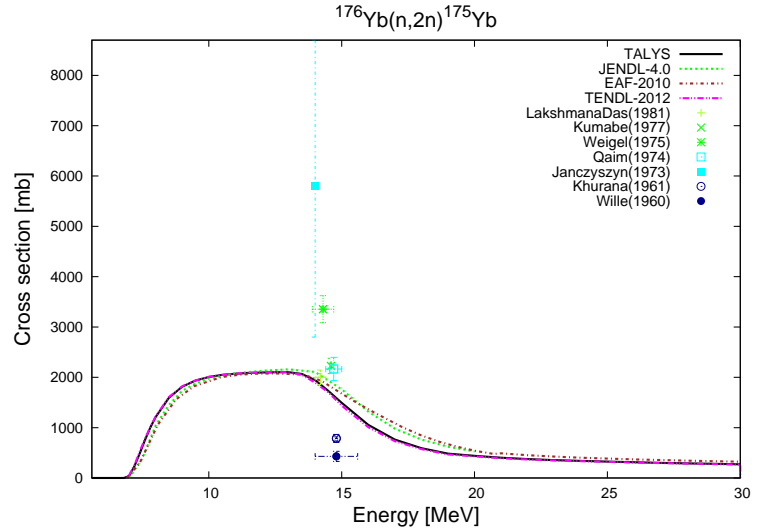
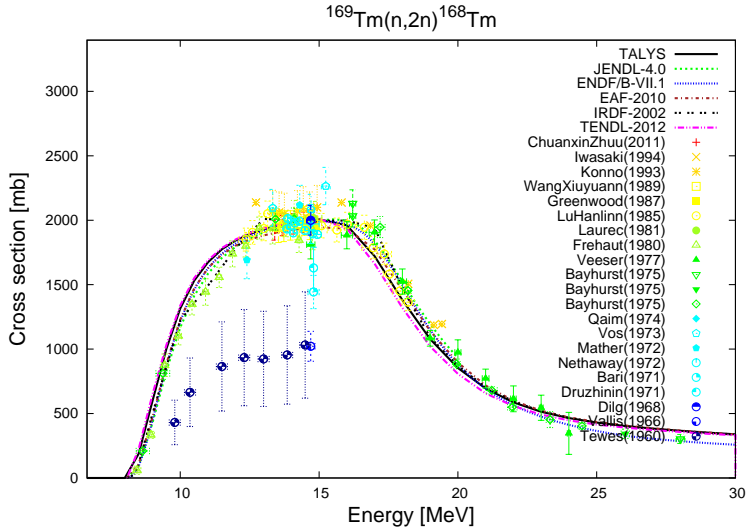
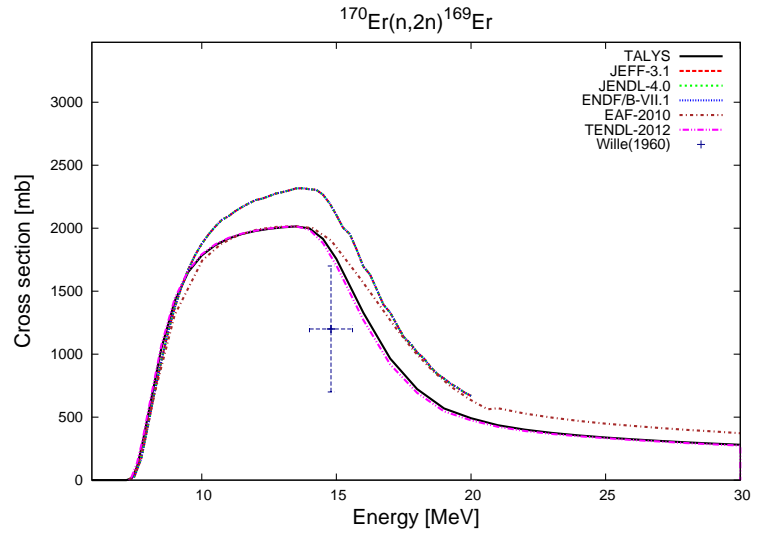
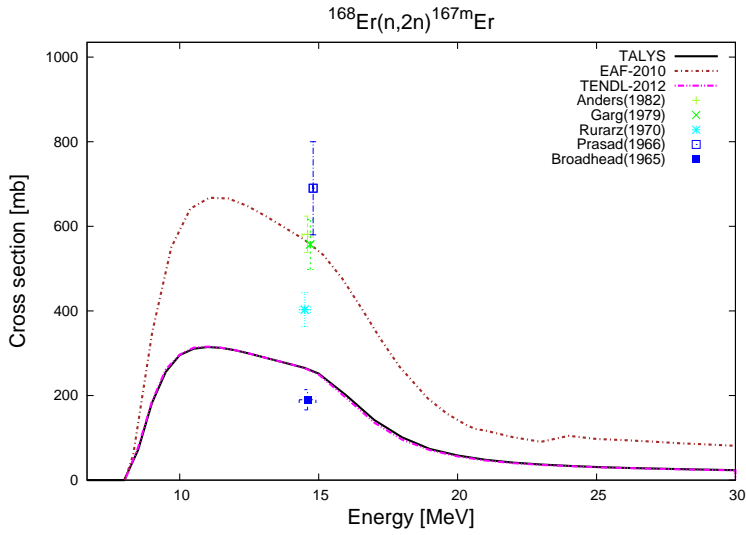
$^{140}\text{Ce}(n,2n)^{139\text{m}}\text{Ce}$  $^{142}\text{Ce}(n,2n)^{141}\text{Ce}$  $^{141}\text{Pr}(n,2n)^{140}\text{Pr}$  $^{142}\text{Nd}(n,2n)^{141}\text{Nd}$  $^{142}\text{Nd}(n,2n)^{141\text{g}}\text{Nd}$  $^{142}\text{Nd}(n,2n)^{141\text{m}}\text{Nd}$ 

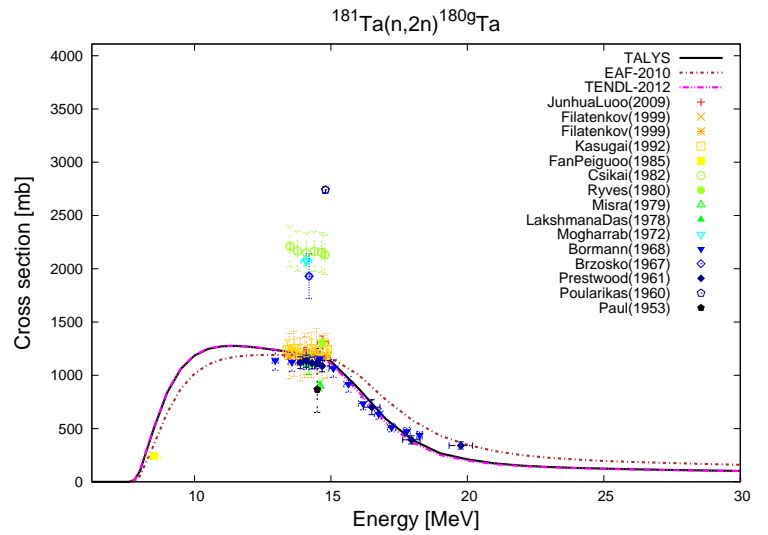
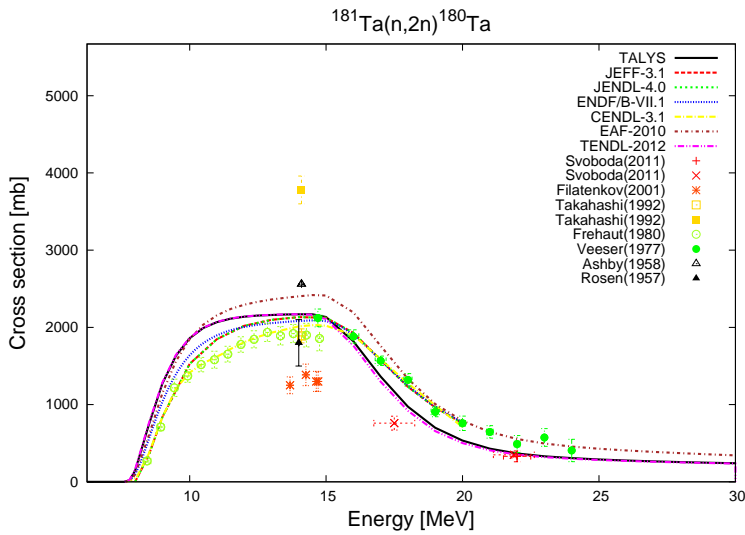
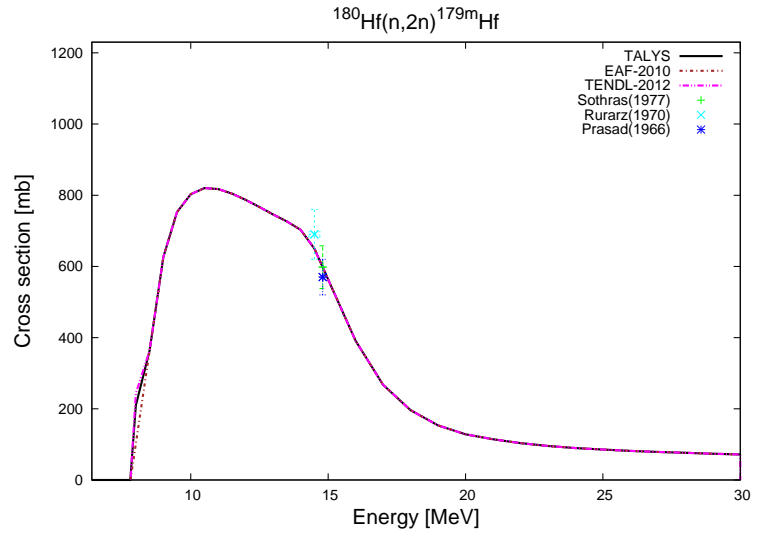
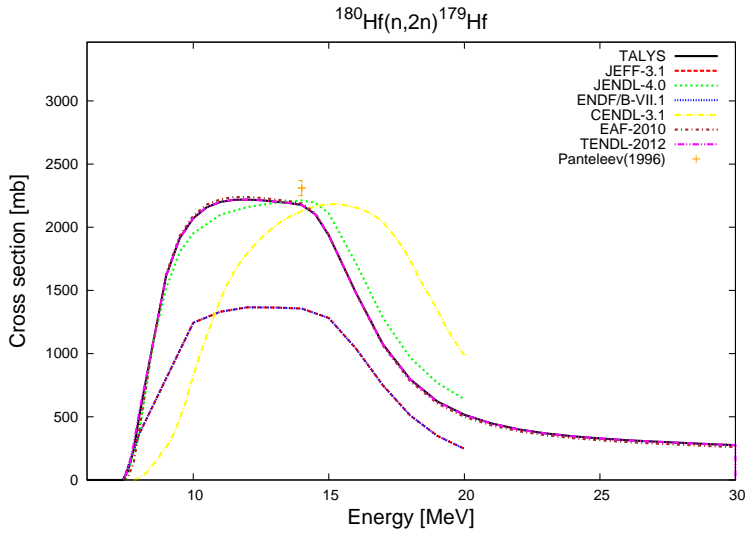
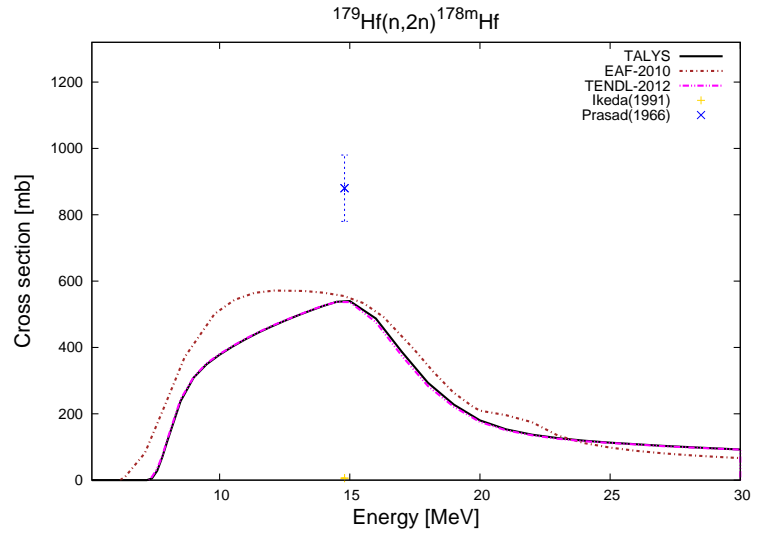
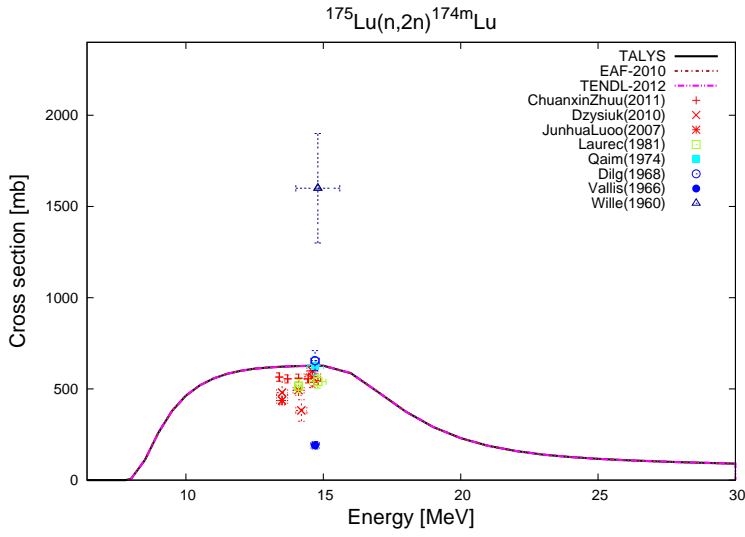


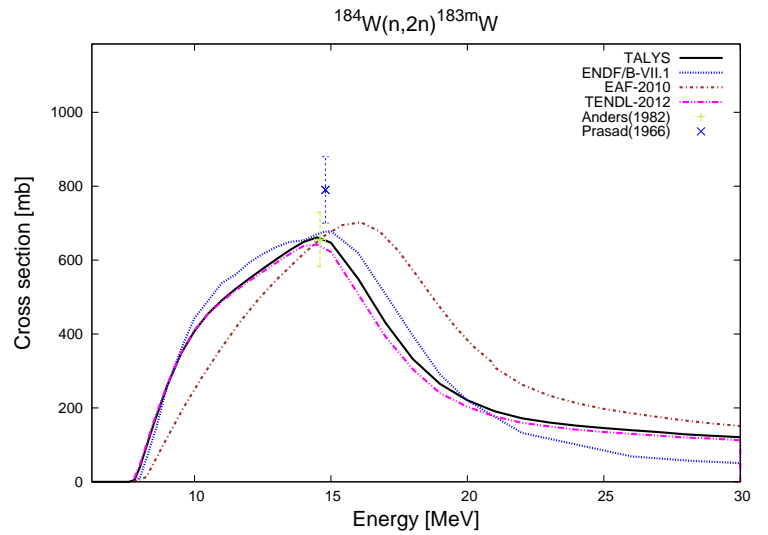
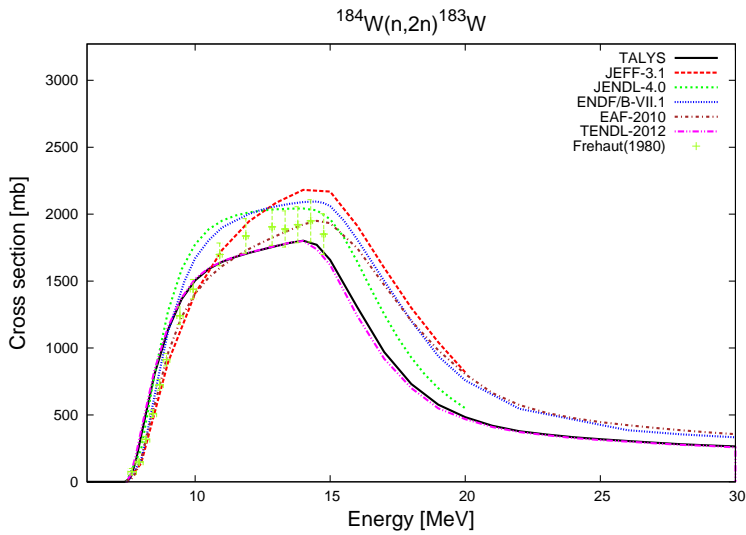
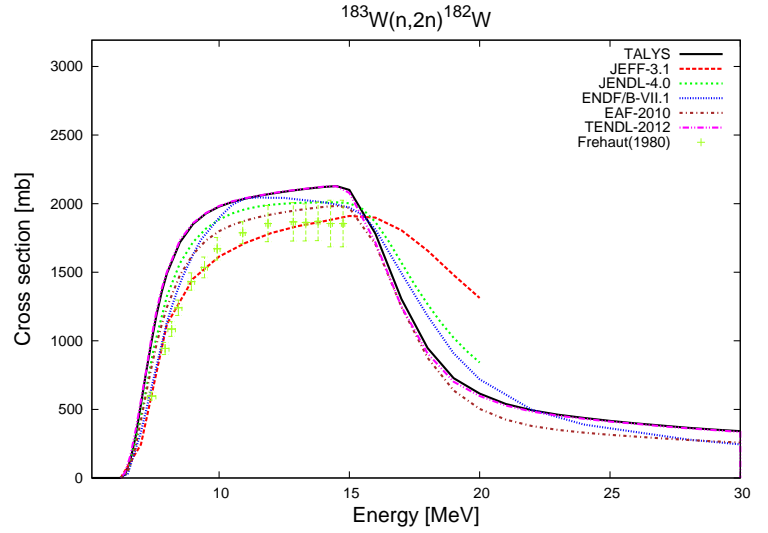
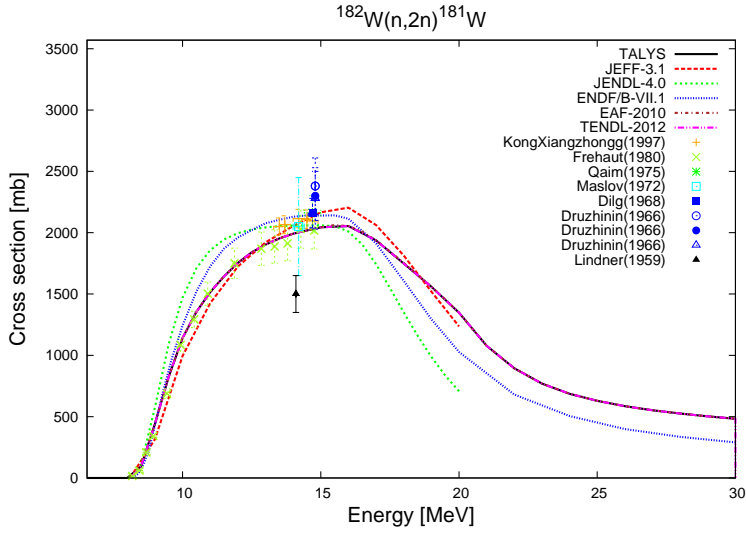
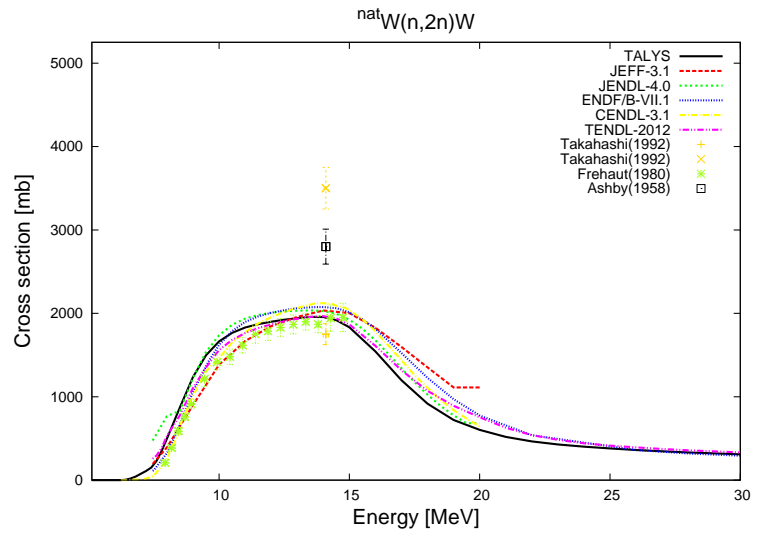
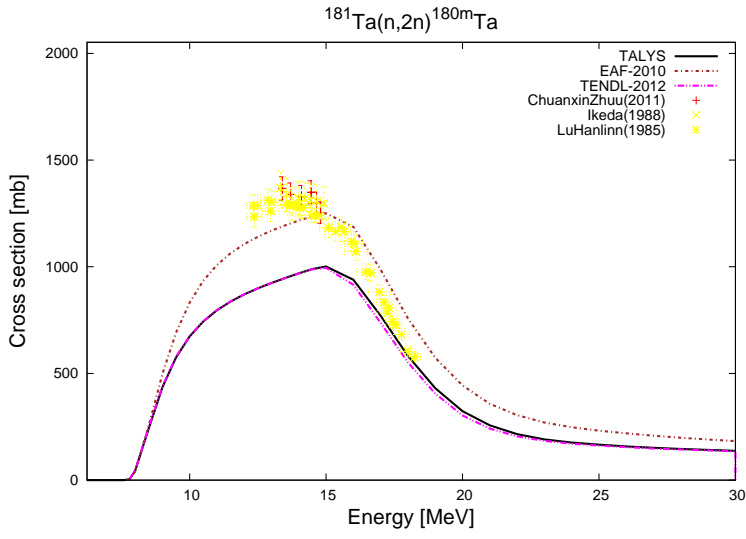


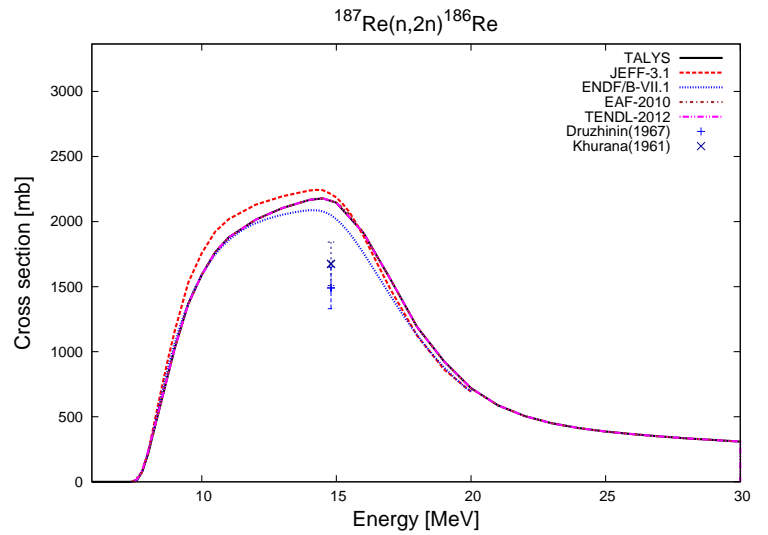
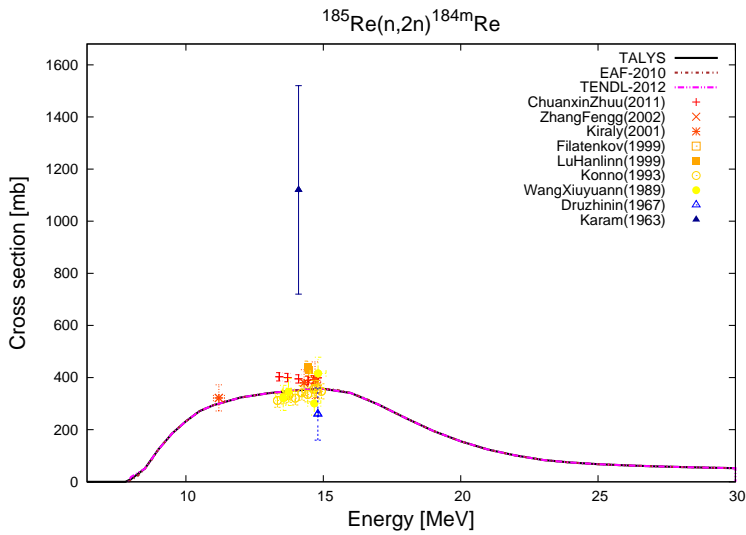
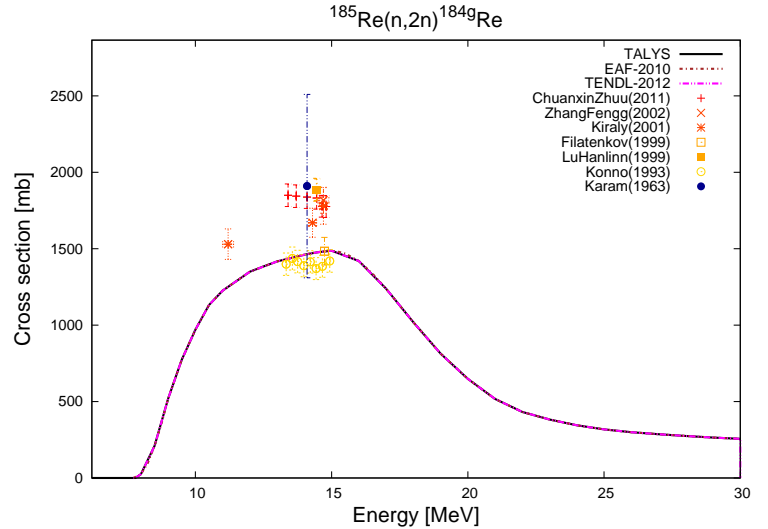
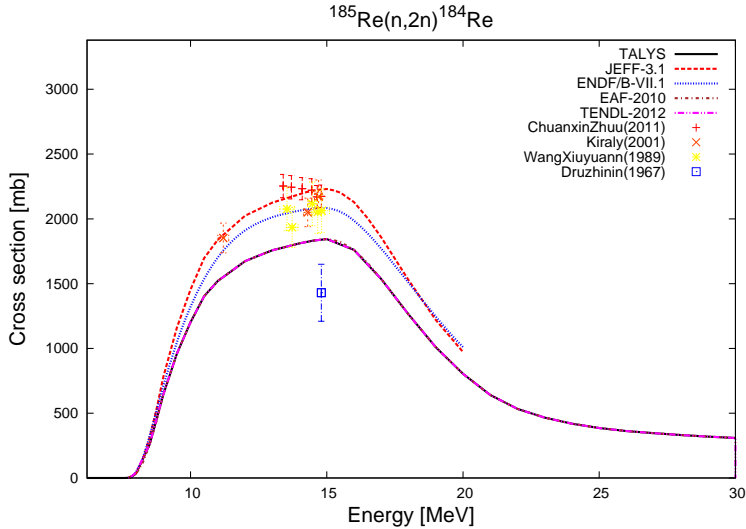
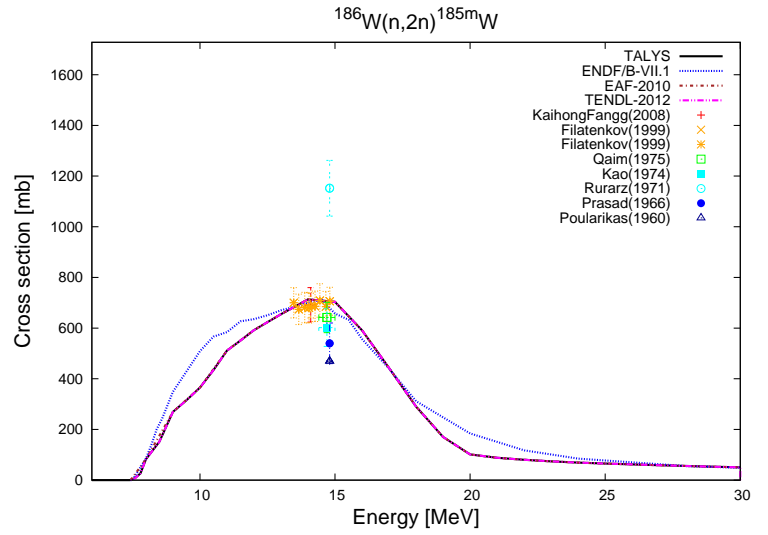
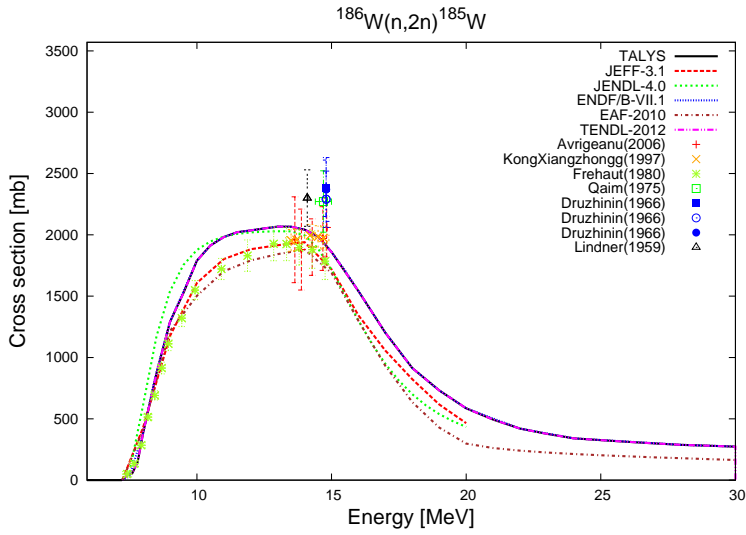


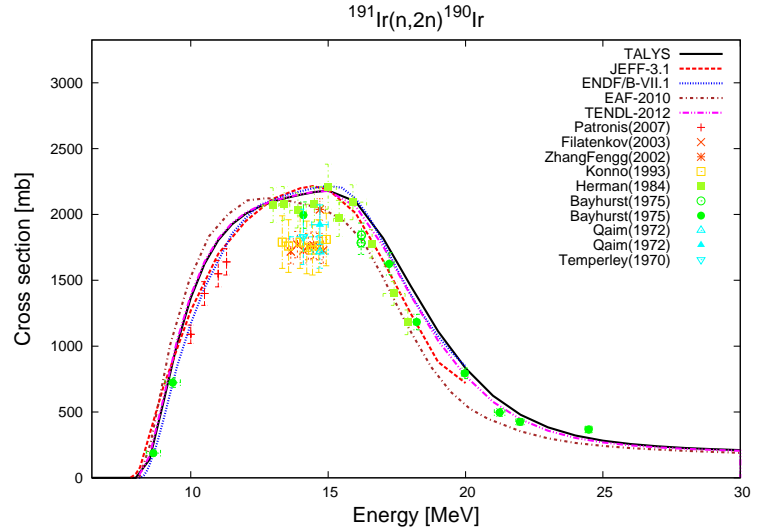
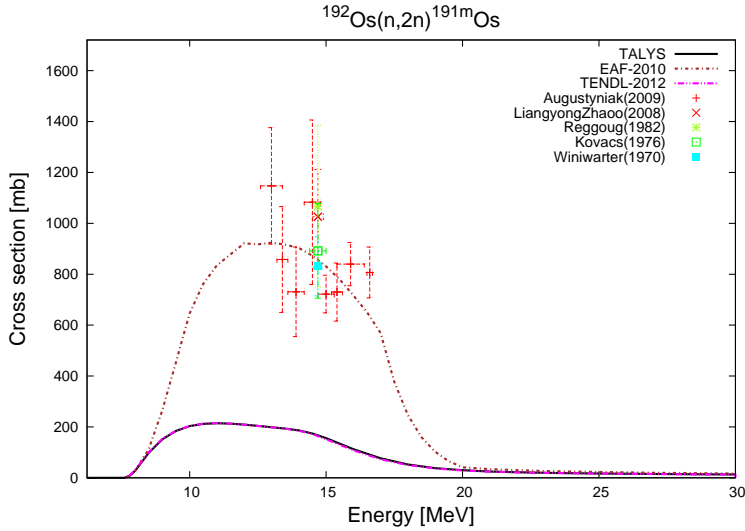
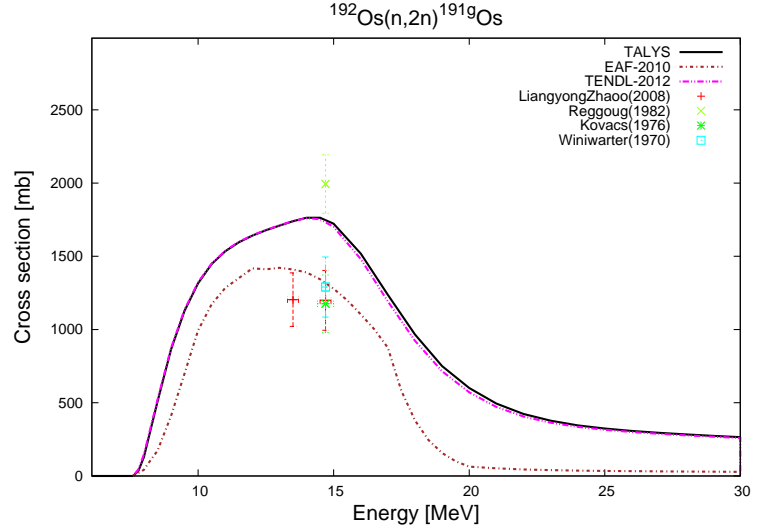
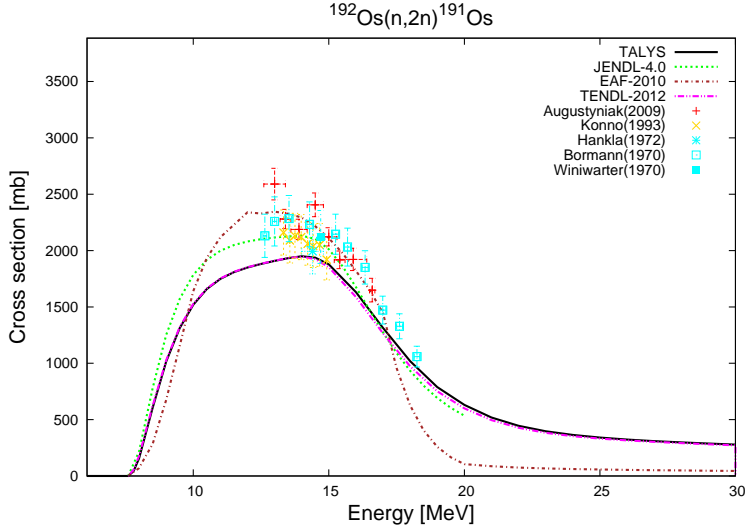
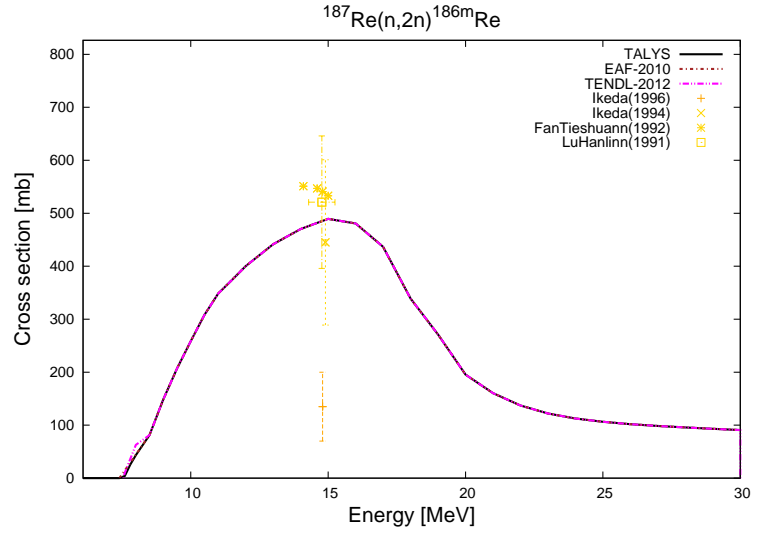
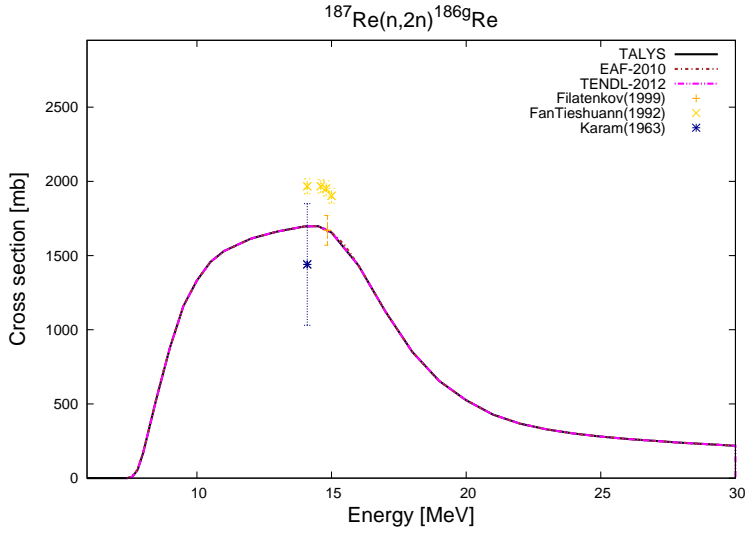


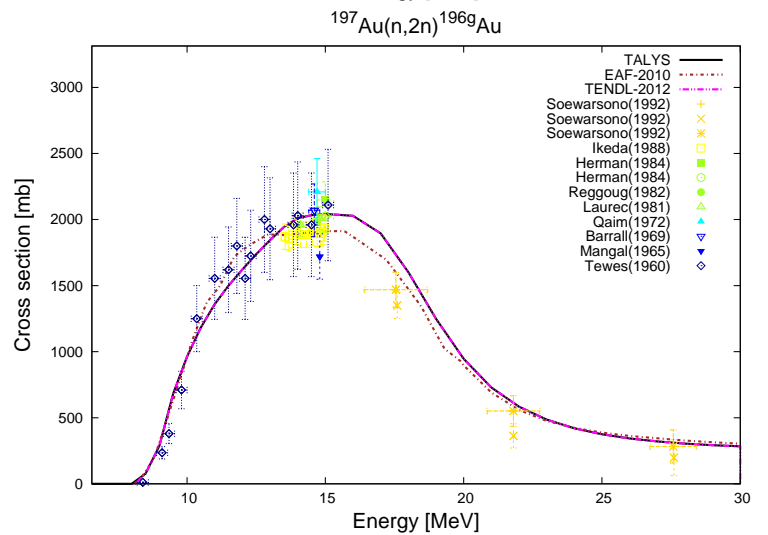
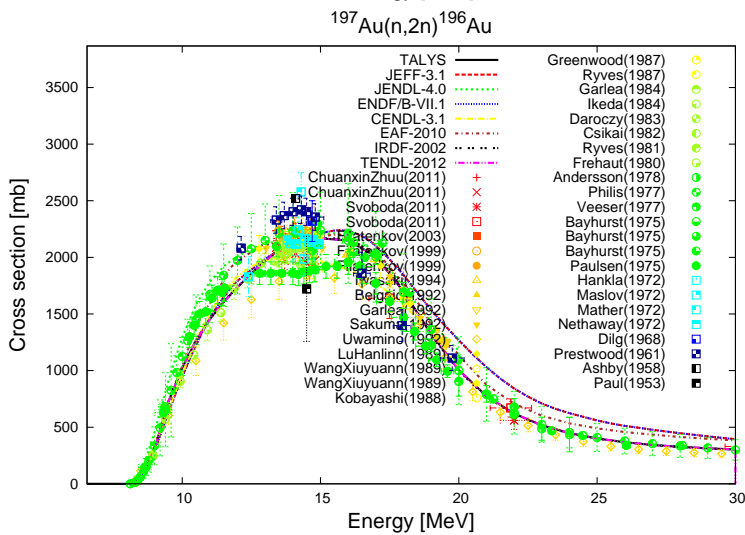
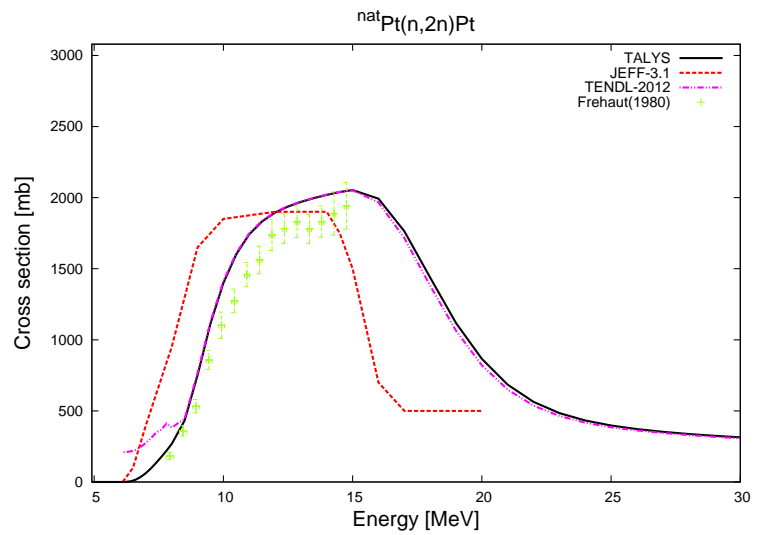
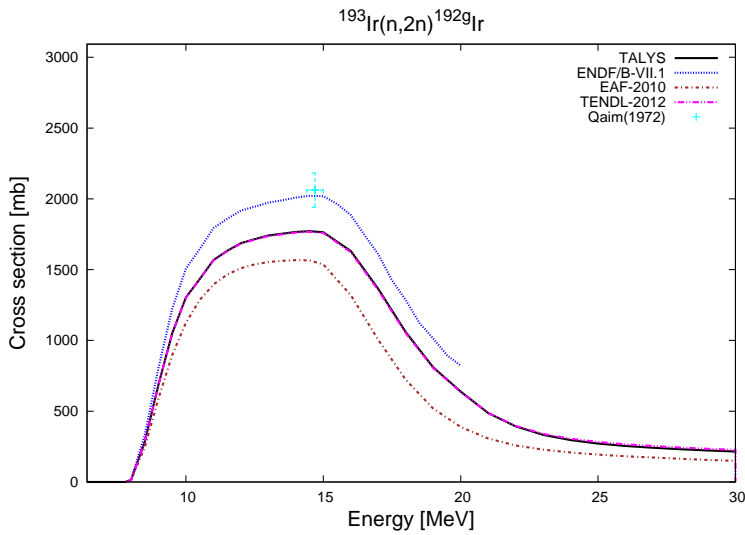
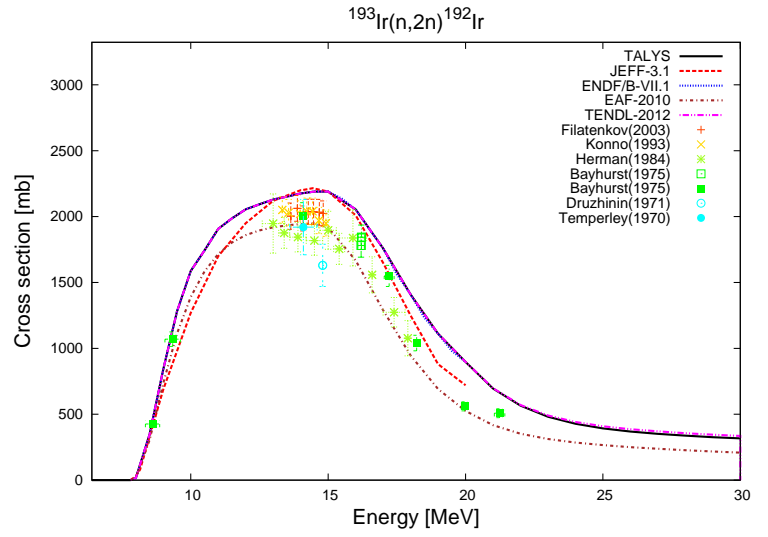
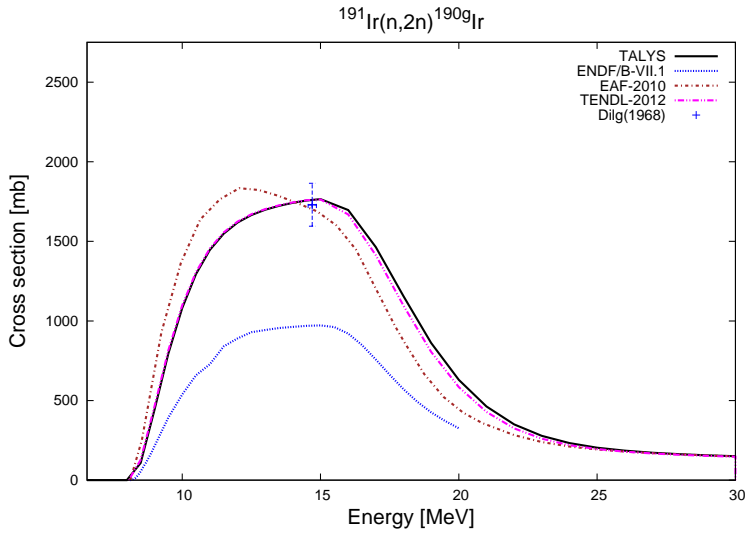


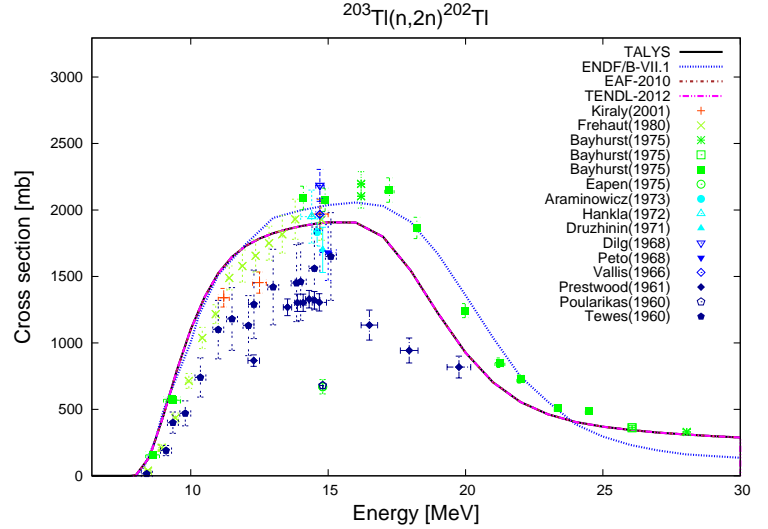
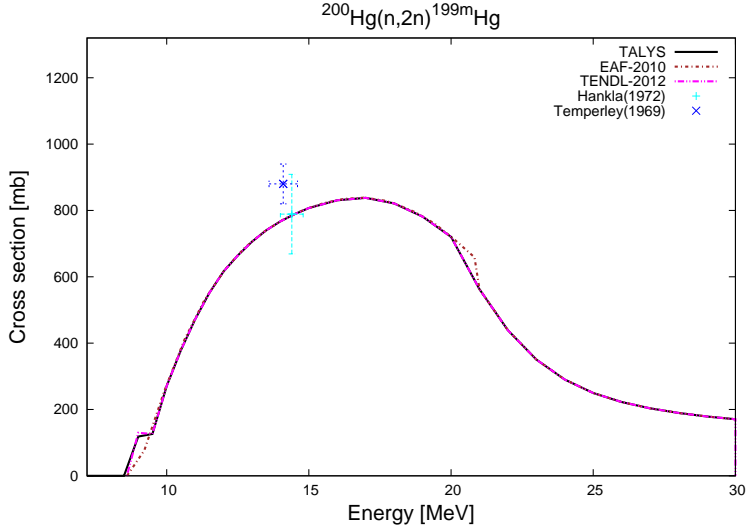
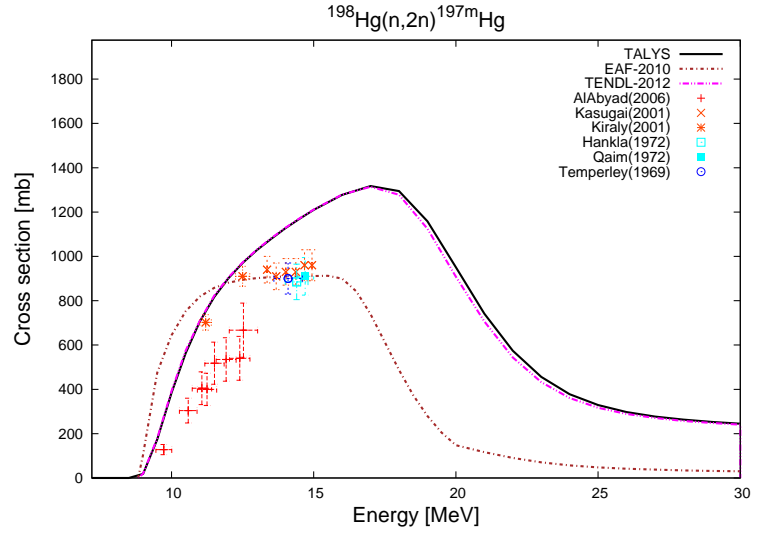
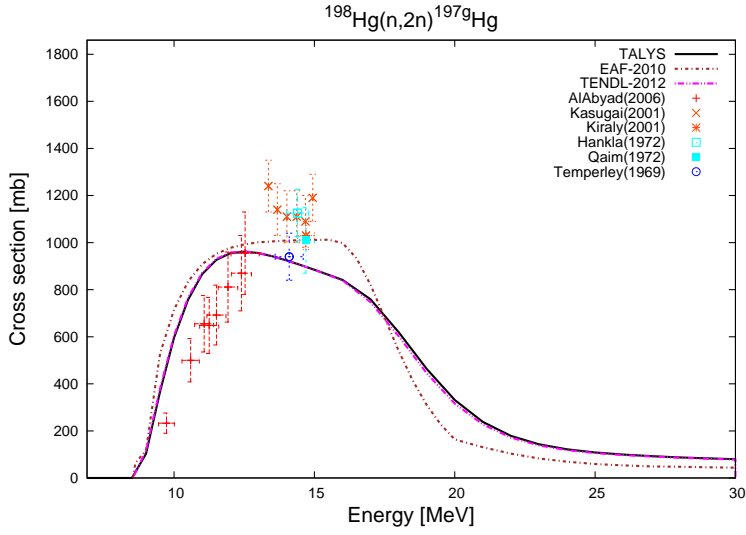
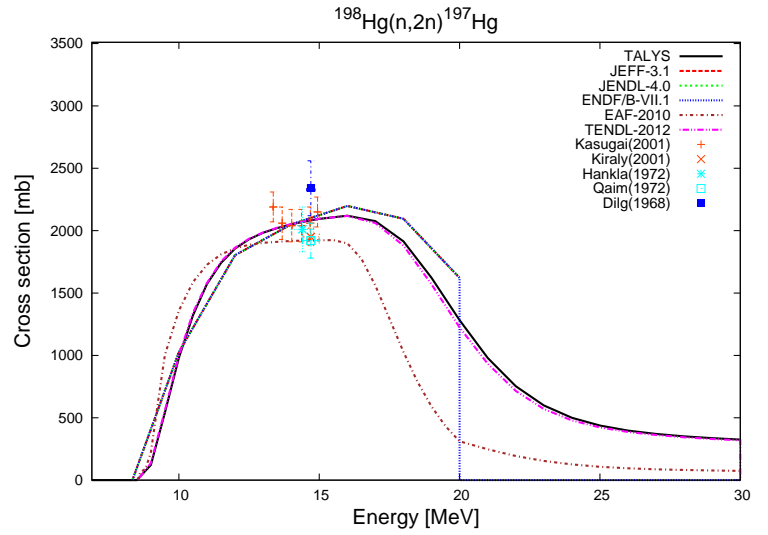
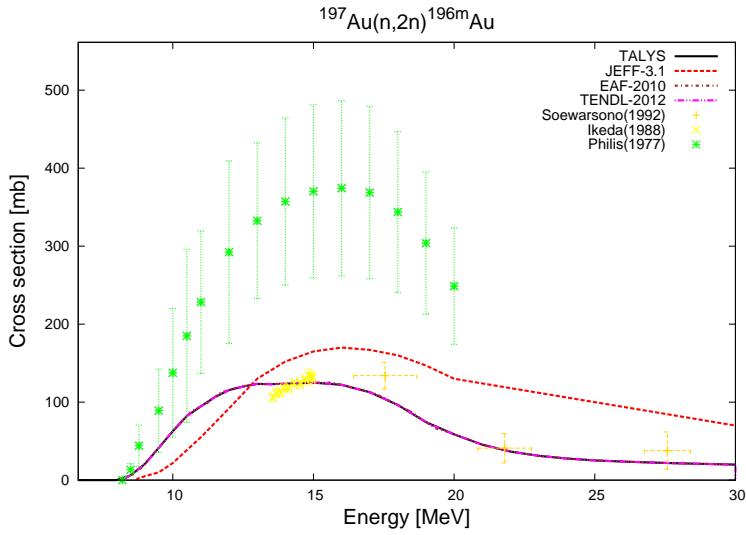


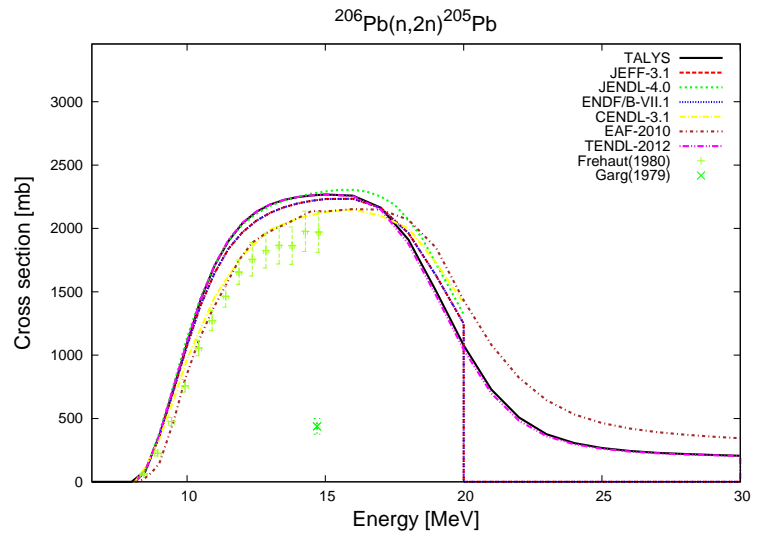
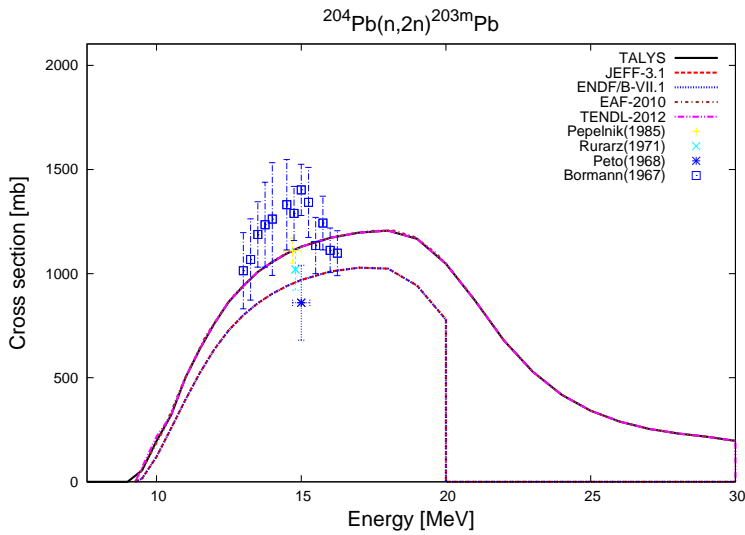
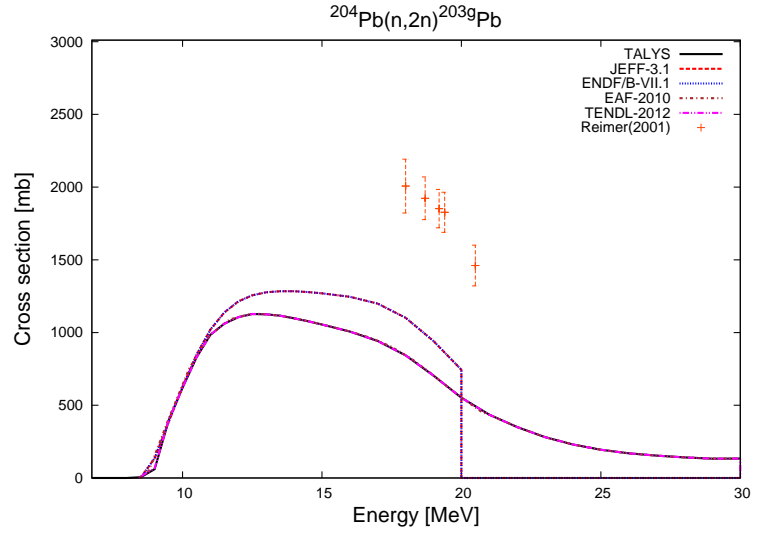
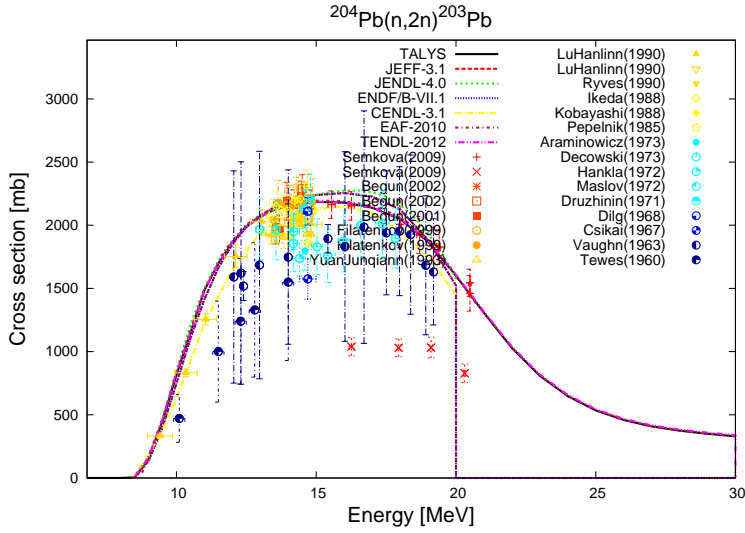
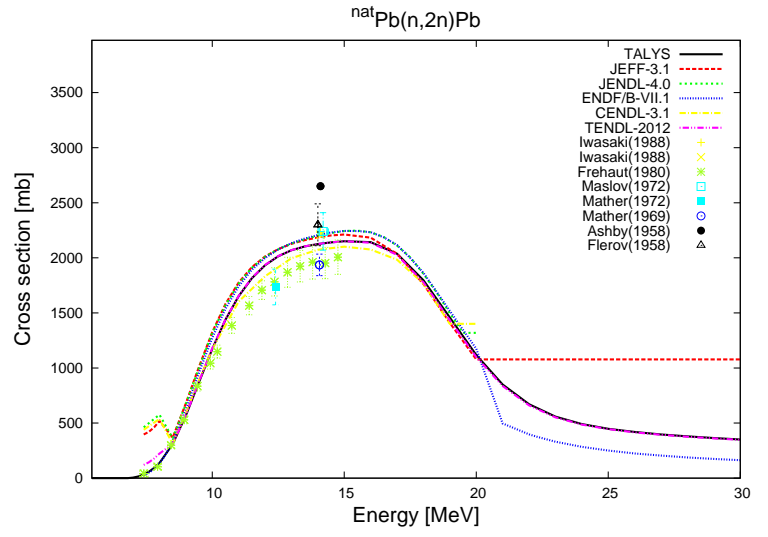
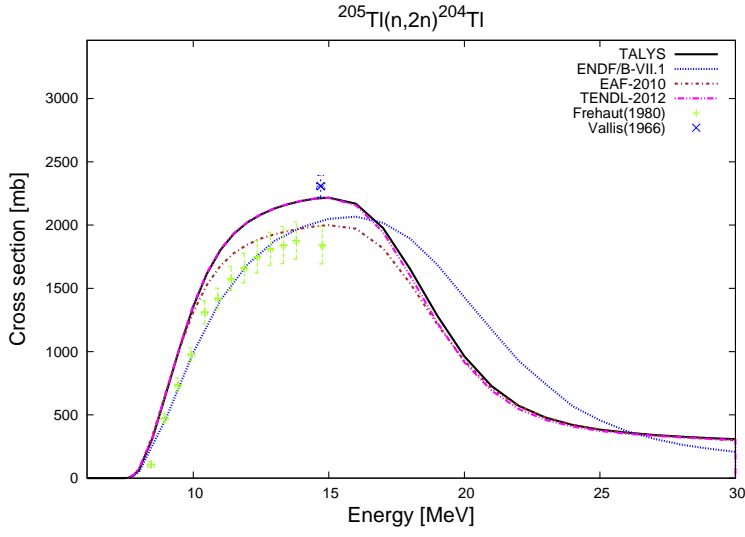


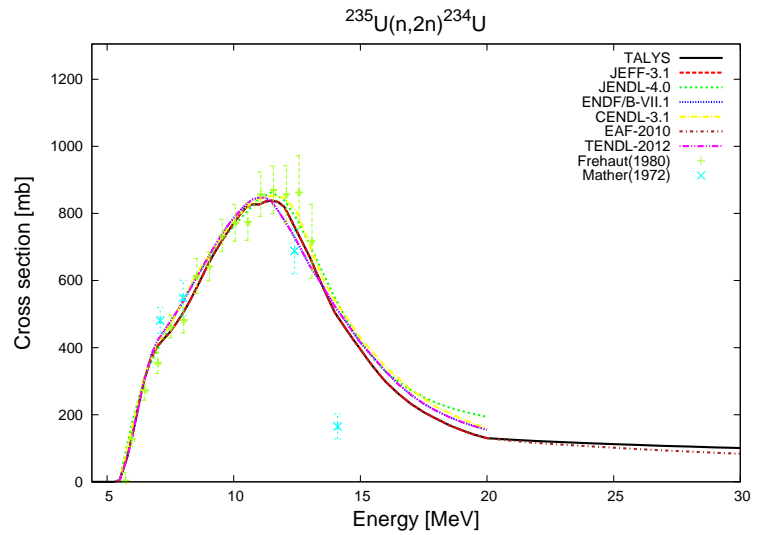
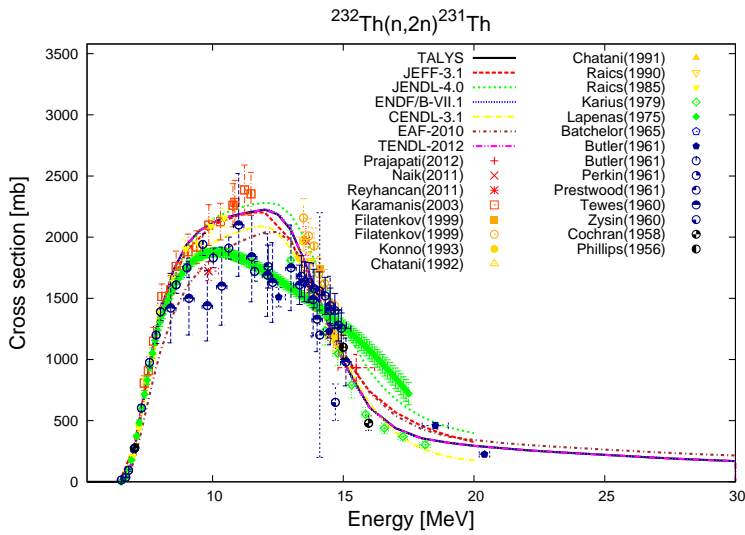
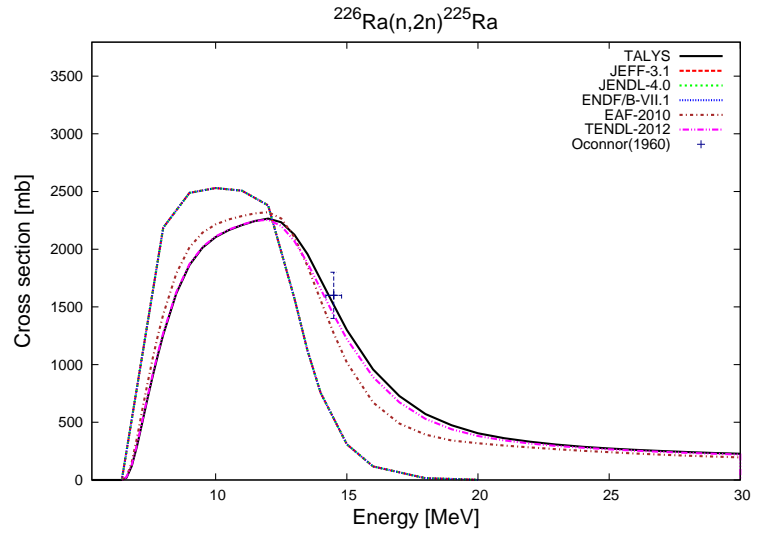
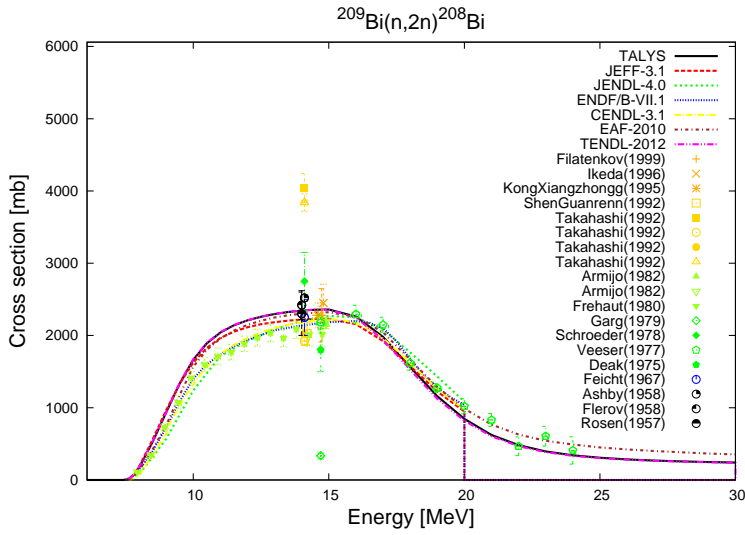
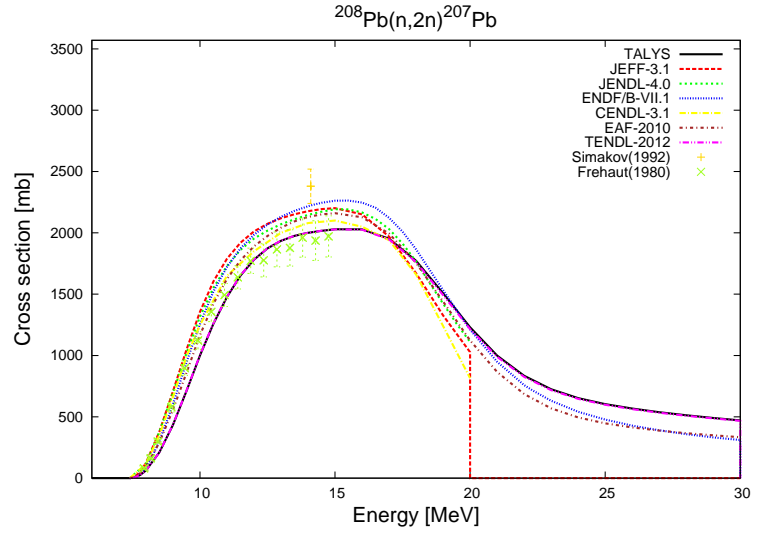
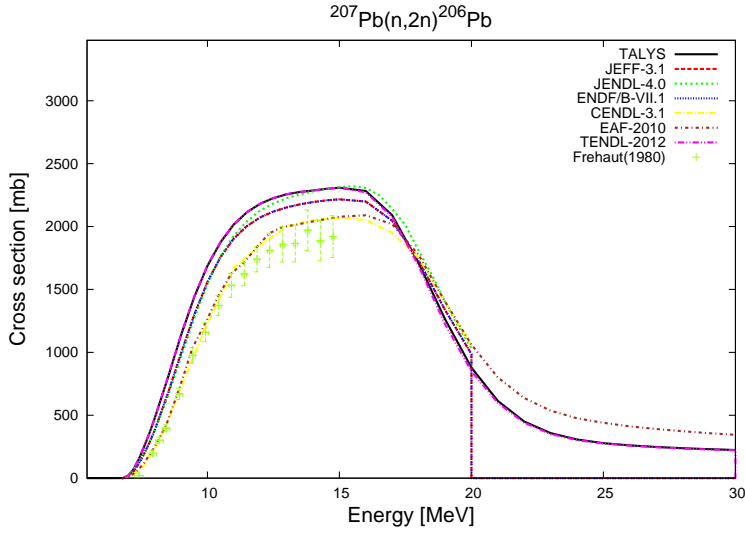


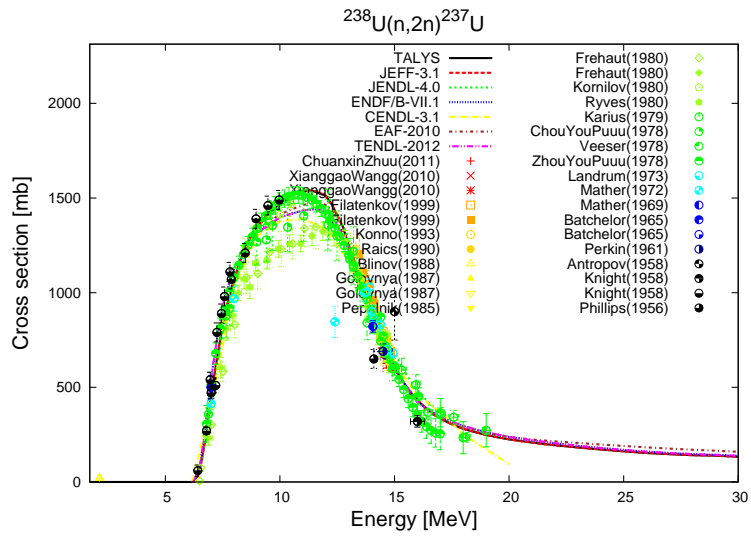


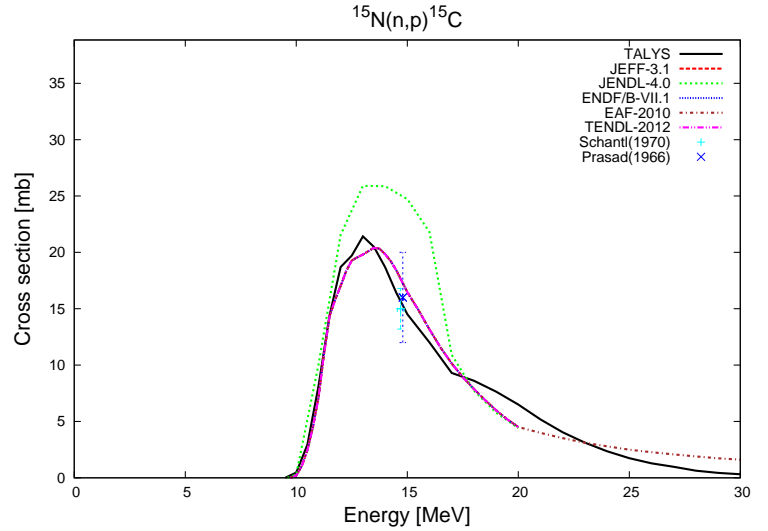
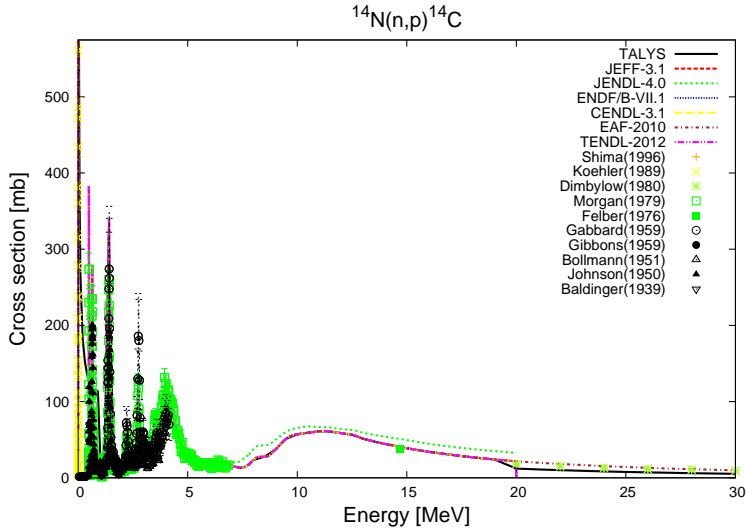
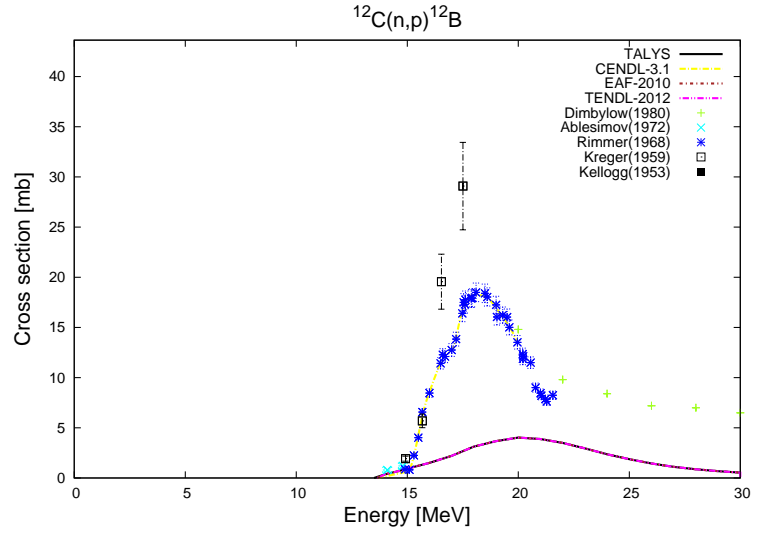
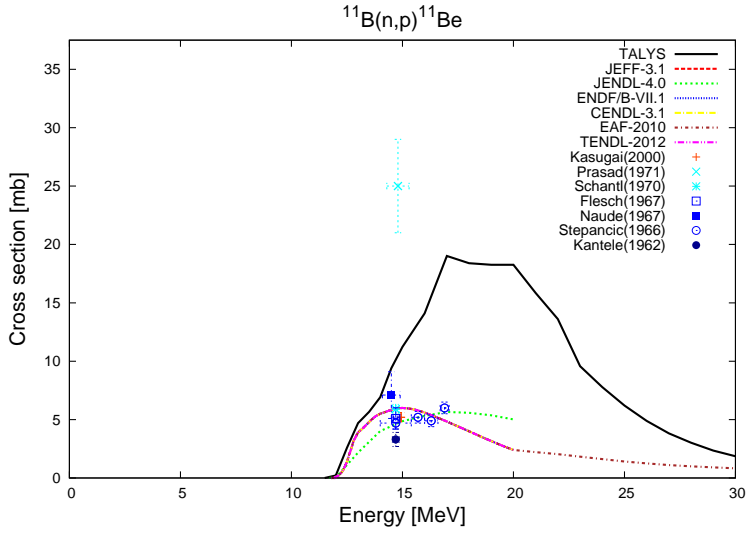
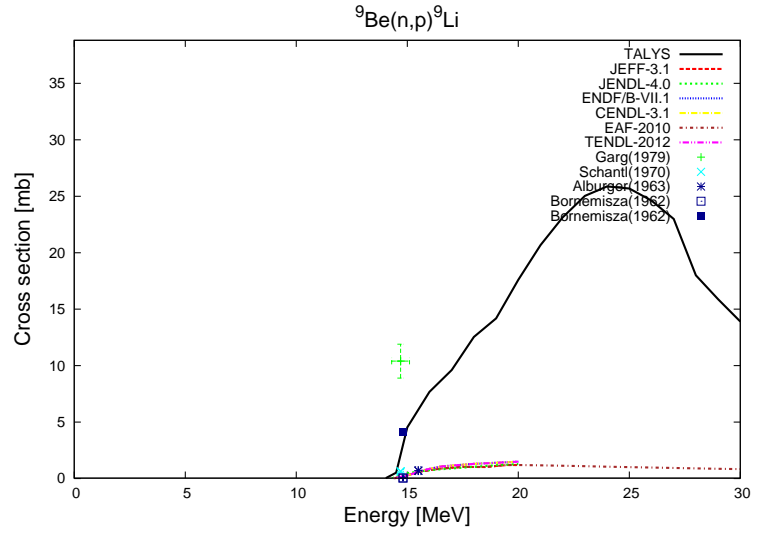
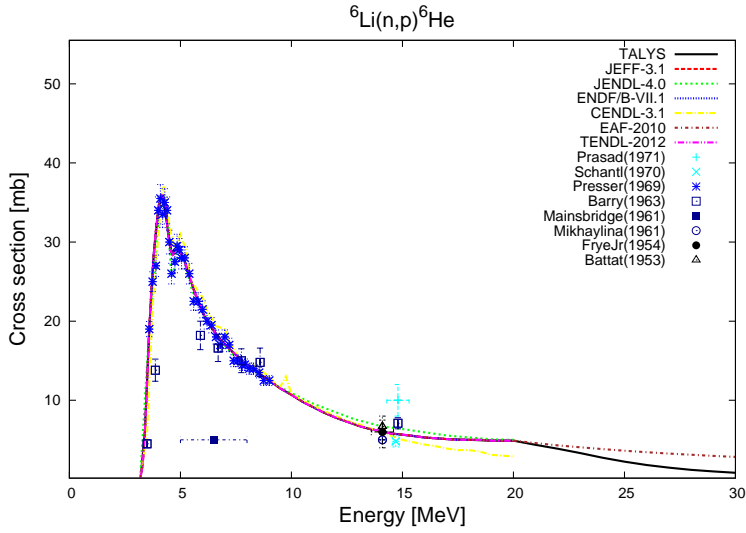


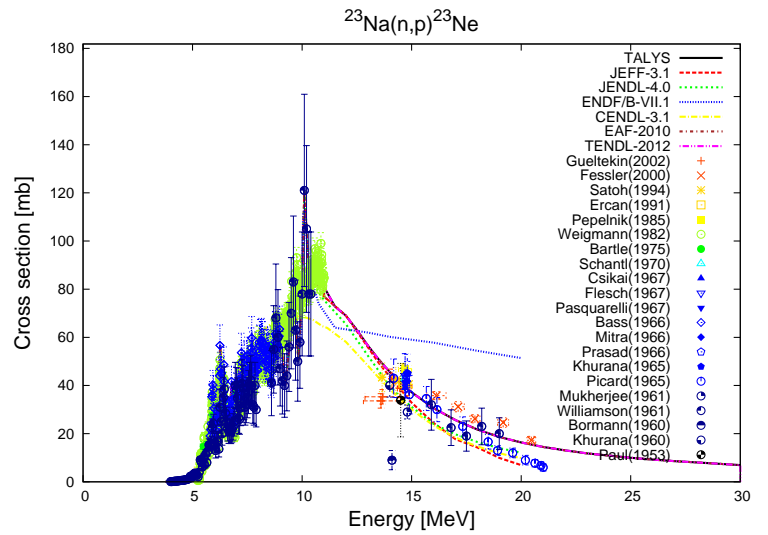
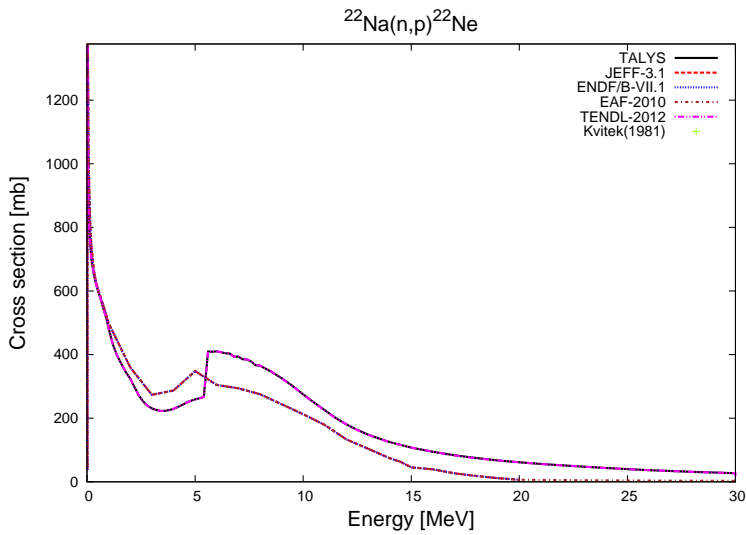
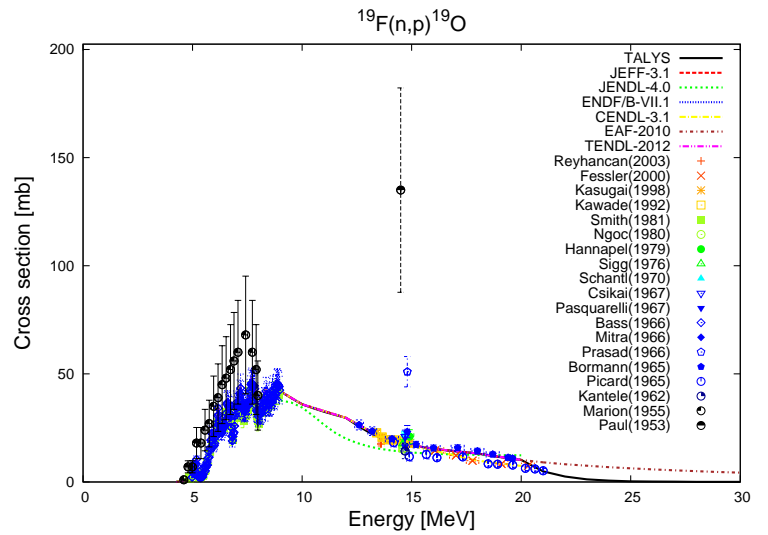
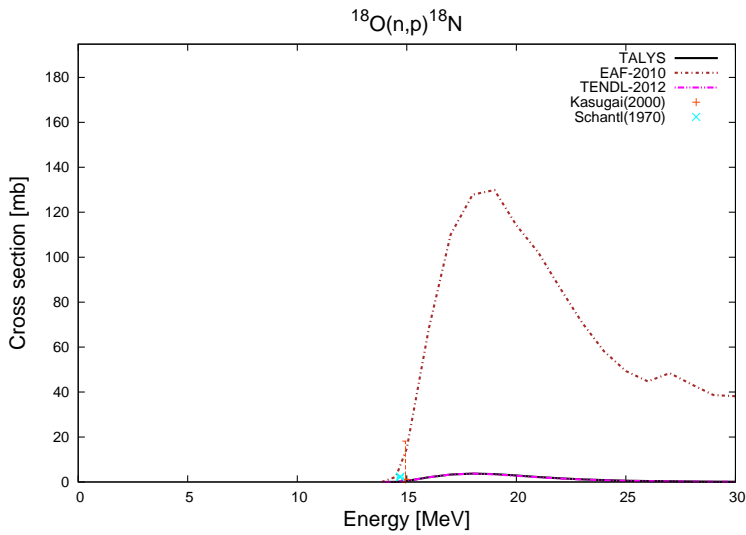
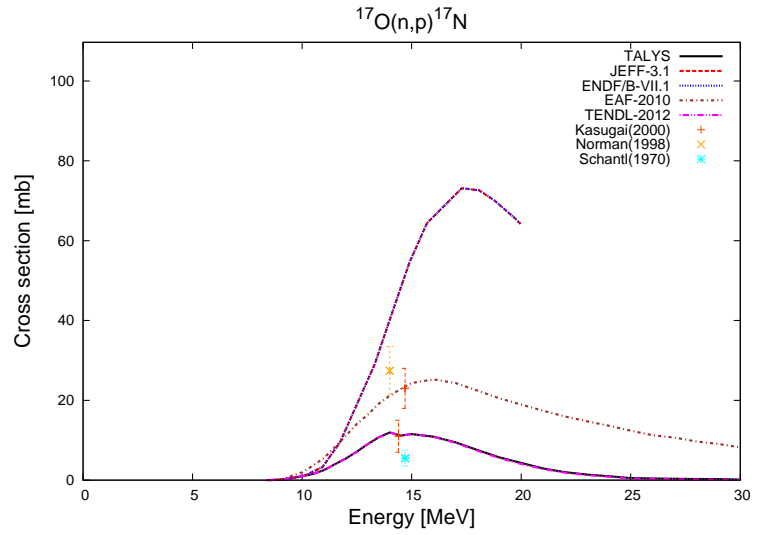
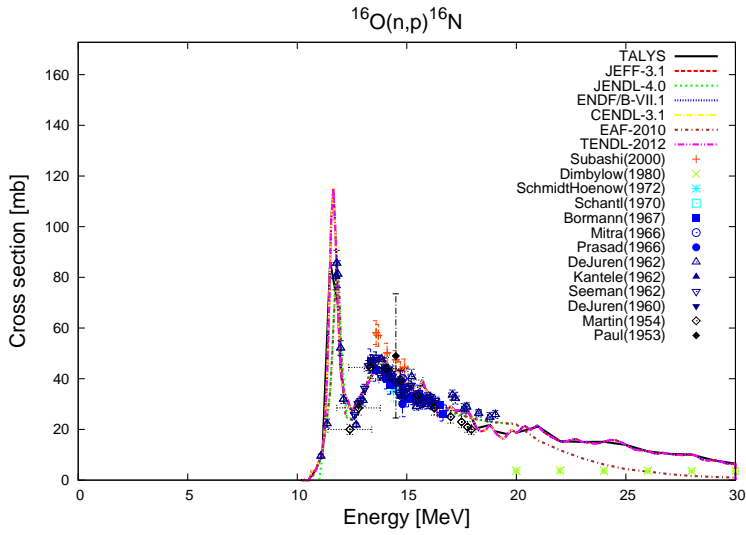


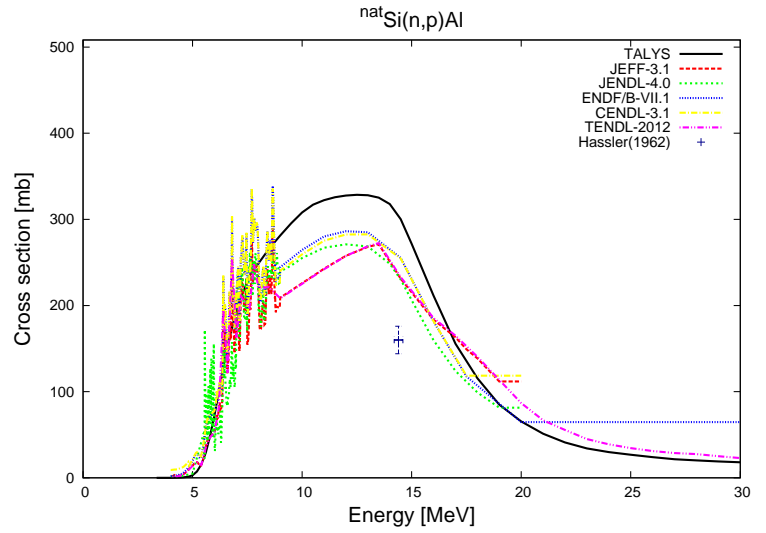
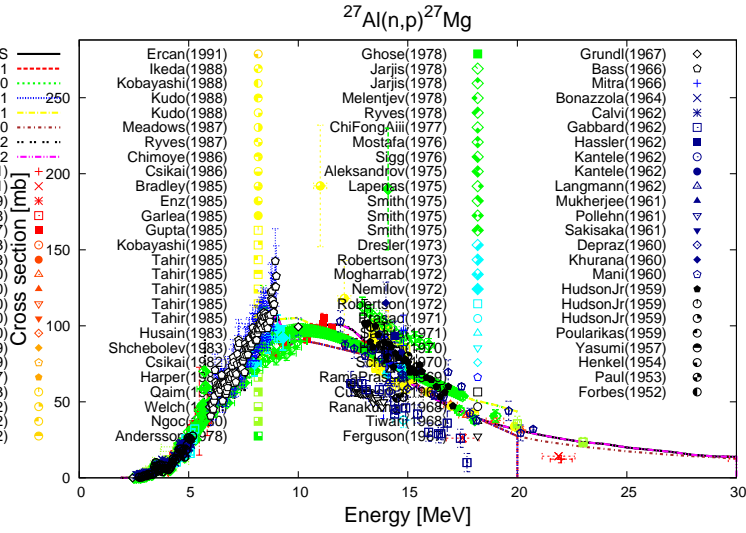
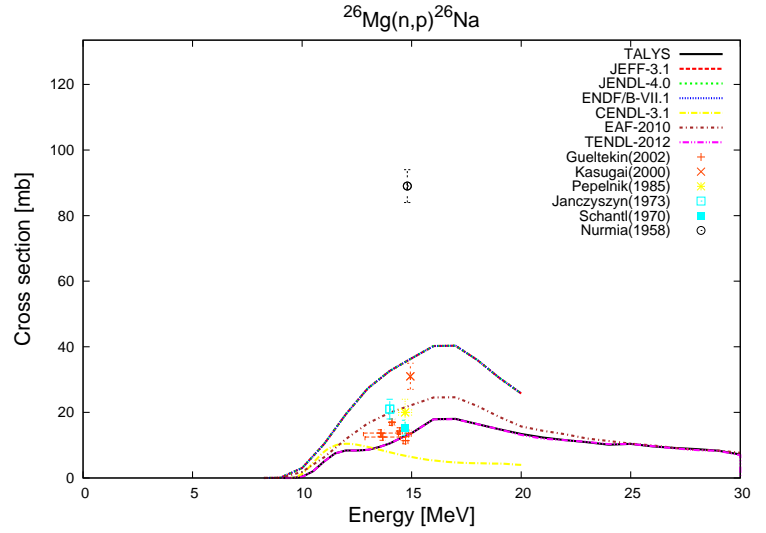
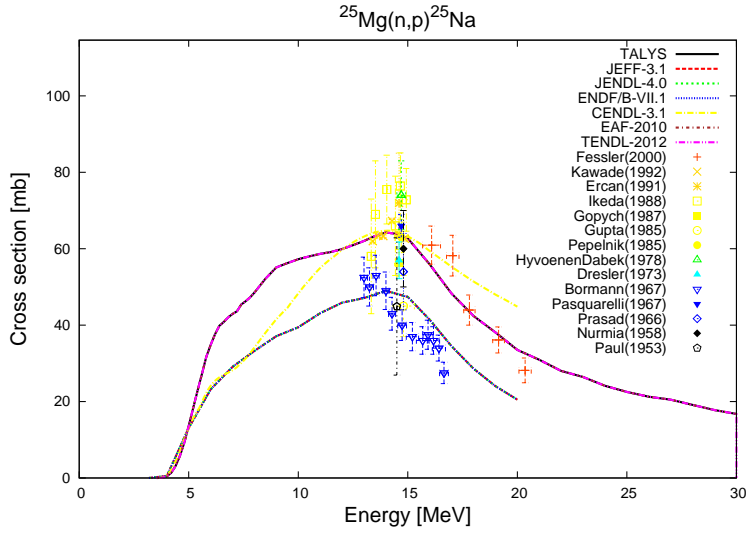
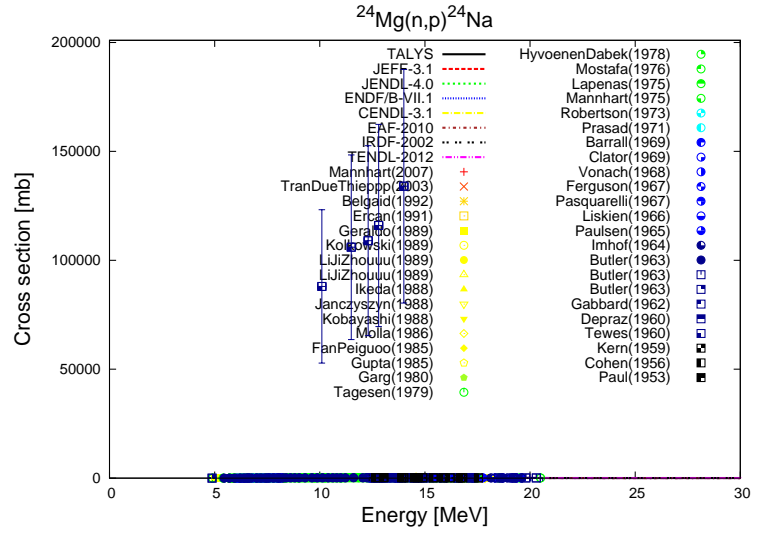
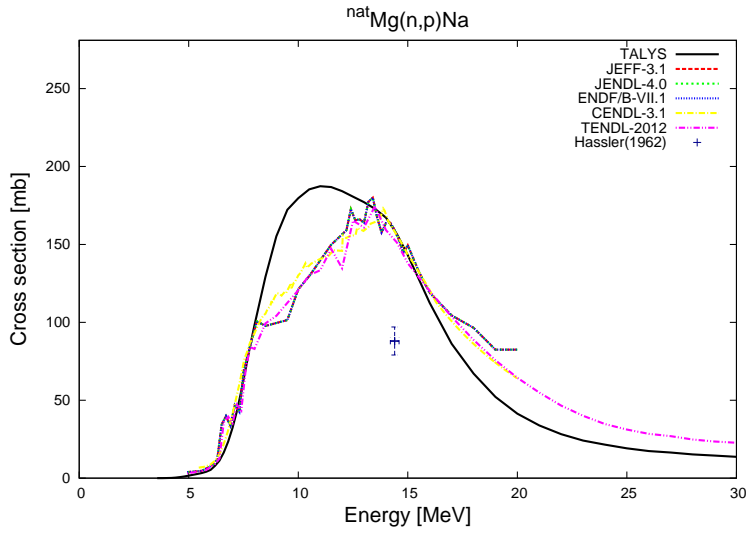


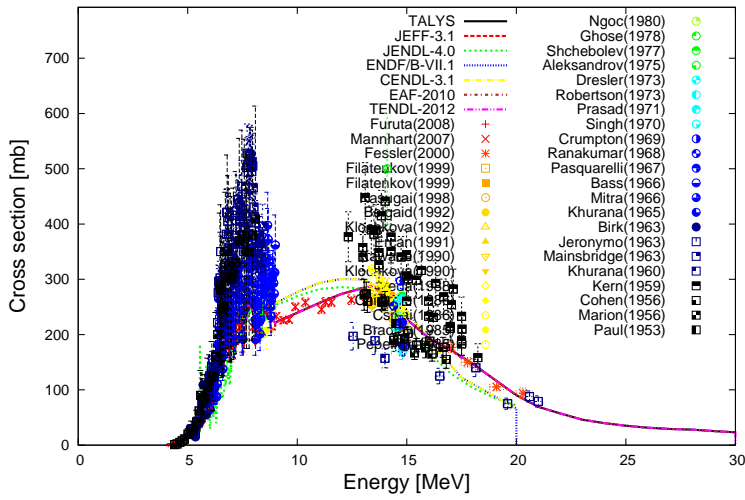
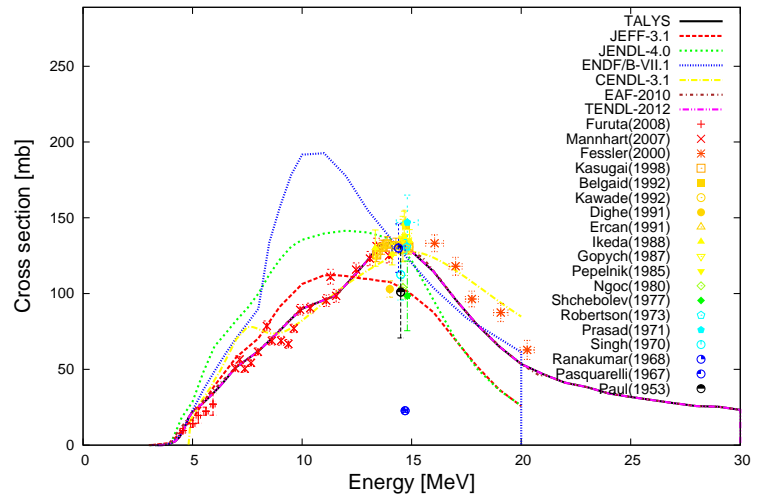
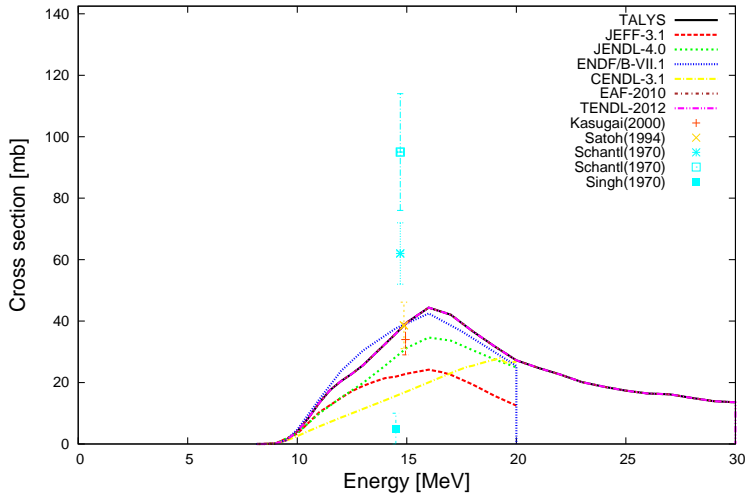
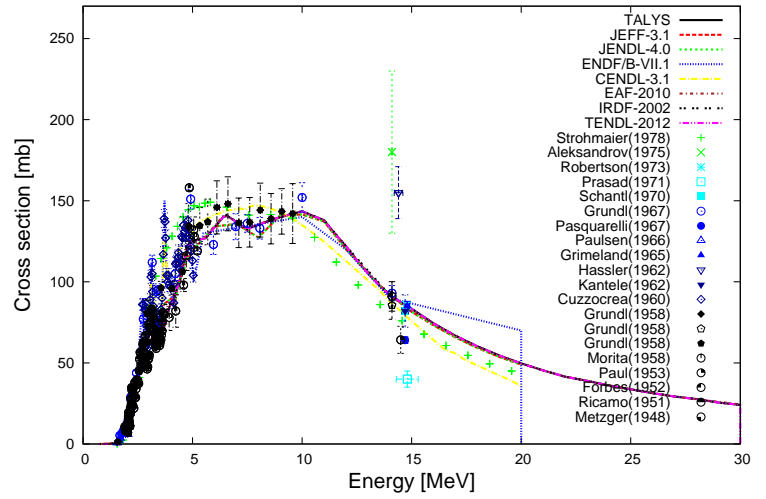
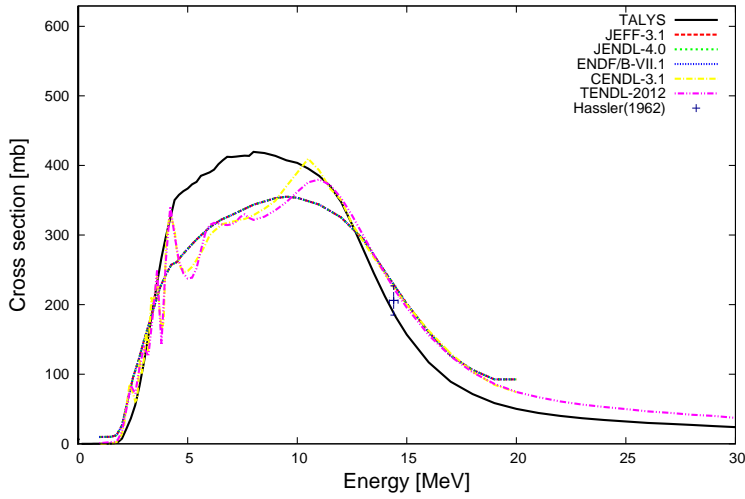










$^{28}\text{Si}(n,p)^{28}\text{Al}$  $^{29}\text{Si}(n,p)^{29}\text{Al}$  $^{30}\text{Si}(n,p)^{30}\text{Al}$  $^{31}\text{P}(n,p)^{31}\text{Si}$  $^{nat}\text{S}(n,p)\text{P}$  $^{32}\text{S}(n,p)^{32}\text{P}$ 