VS1000 application note:

Converting mp3 files to lower bitrate Ogg Vorbis format

	Revision history

Rev.	Date	Author	Affected chapters	Description
1.0	15.11.2007	TK & HH	All	Original version
1.1	16.11.2007	TK	All	Minor corrections
1.2	19.11.2007	TK	TOC	Page numbers for Table of contents regenerated

Table of Contents

1. Introduction

This document presents recommendations for basic parameters when transcoding MP3 files or compressing CD content to Ogg Vorbis format. The purpose is to give to our valuable customers information on how to create high‑quality Ogg Vorbis content.

Before using the information in this document, you should get the free Ogg Vorbis encoder "oggenc", minimum version 1.0.2 (2005), or a version with aoTuV tunings. While the instructions will work with older versions, sound quality will not be as good. Download oggenc at http://www.vorbis.com/.

It is also strongly encouraged to get VorbisGain, a program that tags files with loudness information (Chapter 2.4). For this example, vorbisgain v0.36 has been used. Download VorbisGain at http://www.sjeng.org/vorbisgain.html.

This chapter will introduce the Ogg Vorbis format, as well as the terms variable bitrate and transcoding. Chapter 2 will tell how to actually compress an audio file by using correct parameters for the best sound quality. The examples have been done in Linux operating system environment. Chapter 3 will give step by step instructions for Windows users to use Cdex program at http://cdexos.sourceforge.net/ for the conversion.

2. The Ogg Vorbis Format

Ogg Vorbis is a license‑free audio codec that allows for high‑quality sound around a large gamut of bit‑rates, ranging from roughly 10 to 500 kbit/s.

With a given bit‑rate, Ogg Vorbis is capable of surpassing the quality of MP2, MP3, WMA and AAC LC, particularly in low‑bit‑rate applications.

An Ogg Vorbis file is a Vorbis compressed file that has been encapsulated inside an Ogg stream. The default file extension for Ogg Vorbis files is ".ogg".

An Ogg Vorbis file consists of two major parts: header and actual data. Every Ogg Vorbis file begins with a header which contains compression tables that the decoder needs to correctly decode actual audio data. Only after the header has been fully received, Vorbis audio may be decoded. The size of the header information is around 4 KiB for most encoders. However, for VLSI Solution's VS1053 circuit encoder, header size is roughly 1.5 KiB.

After the header has been decoded, full random access is available in an Ogg Vorbis file. Unlike MP3 format, an Ogg Vorbis file contains its own time code, so absolute position in a file can be displayed after a random access operation like fast forward / reverse.

Because the header is relatively big (4 KiB), Ogg Vorbis is not particularly well suited for encoding very short, low bit‑rate samples. E.g. a 16 kbit/s audio sample that lasts for one second would take 2 KiB for the actual audio data, so the default header triples the size of the file. However, for normal applications where audio lasts for more than 10 seconds, this header overhead is negligible.

2.1. Variable Bit‑Rate

For a long time almost all MP3 files had a constant bit‑rate, typically 128 kbit/s. While constant bit‑rate (CBR) makes it easy to calculate a relation between file size/position and song length, and while it makes it easier to stream files, CBR is not optimal from a quality point of view. When there is a difficult passage in music, there may not be enough bits available to encode audio flawlessly. Conversely, in passages that are easier to encode there may be bits left unused because the encoder didn't need all the bits available in the bitstream.

Where CBR tries to keep the bit‑rate constant regardless of content, Variable Bit‑Rate (VBR) tries to keep quality constant. This will have the effect of changing the bit‑rate of the file depending on how difficult a particular portion of music is to encode. Because sound quality is always kept constant, VBR has a bit advantage to CBR. Thus, VBR allows either for smaller files with the same sound quality or equally sized files with better sound quality.

Ogg Vorbis is by nature a variable bit‑rate format and works best when this feature is not limited. Depending on the material the bit‑rate can have significant fluctuations: e.g. songs that have been compressed with a quality setting with a nominal bit‑rate of 96 kbit/s can easily have average bit‑rates between 80 to 105 kbit/s, or even more with extreme material.

2.2. Transcoding

Transcoding is the process of converting from one lossy format to another, like from MP3 to Ogg Vorbis.

It is important to understand that although Ogg Vorbis is a better format than MP3, **transcoding never enhances sound quality**. What has been lost in the initial MP3 compression stage, will stay lost when converting to Ogg Vorbis, and there will always be some additional quality loss. While the extra quality loss can be small enough that it may be impossible to tell, the end result still is never better than the original file was.

If possible, it it recommended that users should avoid transcoding to avoid generation loss. If transcoding is performed (e.g. to get smaller files for a portable device), it is a good idea to keep the original files.

For best results, compress Ogg Vorbis files from uncompressed sources, like CDs.

3. Compressing

The whole process of getting Ogg Vorbis files looks roughly like this:

- Preparation: Decode an MP3 / WMA / AAC etc file or rip a CD to WAV format and make sure that you have "oggenc" (Chapter 3.1).
- Compressing to Ogg Vorbis (Chapters 3.2 and 3.3).
- Optional post‑processing with VorbisGain (Chapter 3.4).

3.1. Preparation

If you intend to transcode (=recompress) MP3 files, you first have to convert them to 16‑bit RIFF WAV format (usually known for a .wav file name extension). However, before you do that, please have a look at Chapters 3.2 and 3.3 for details on transcoding.

If you rip CDs, you have to make sure they are ripped initially into WAV format. It is recommended to use CD sources for compression if possible.

When you have audio in WAV format, you can use `oggenc` to compress files, either in stereo or mono.

Before continuing to the next stage, make sure you have the free Ogg Vorbis encoder *oggenc*, minimum version 1.0.2 (2005).

3.2. Compressing CD Quality Music

If we assume the input file is called file.wav and the output is to be called file.ogg, the following parameters may be used.

It is assumed that the input audio is sampled at the CD 44.1 kHz sample rate. While these commands will also work for other sample rates, quality and file sizes may differ.

For each quality setting, the table shows the command needed to convert the file, a typical bit‑rate for that quality, how many hours of music could be stored to a 1 gibibyte memory card, and a short description of the mode. For reference, also an uncompressed CD bitstream is shown in the table.

CD Quality Compression			
Command line	Typ. kbit/s	Hours / 1 GiB	Description
`oggenc -q -1 file.wav`	45	53	`Smallest files size, better than 96 kbit/s MP3`
`oggenc -q 0 file.wav`	64	37	`Ok quality, equals 112 kbit/s MP3`
`oggenc -q 2 file.wav`	96	25	`Better than most 128 kbit/s MP3 files`
`oggenc -q 4 file.wav`	128	19	`CD transparent in most cases`
`oggenc -q 6 file.wav`	192	12	`CD quality almost always`
`CD`	1411	1.7	`Uncompressed CD`

Graphical frontends usually contain one or several CBR bit‑rate slider(s) and a quality slider that usually goes from ‑1 to 10. CBR sliders should be deactivated, and the quality slider should be set to the "‑q" value presented in the table.

3.3. Compressing Speech Quality Music

Speech requires different coding parameters from music. First, speech can usually be coded in mono, so all the following examples convert audio to mono. Also, to get smaller files, bandwidth

of speech can usually be suppressed without adversely affecting intelligibility. Thus, substantially lower bitrates can be achieved when compared to full‑bandwidth stereo music.

The following table shows the command needed to convert a speech file called file.wav, a typical bit‑rate for that quality, how many hours of music could be stored to a 1 gibibyte memory card, and a short description of the mode.

Speech Quality Compression			
Command line	Typ. kbit/s	Hours / 1 GiB	Description
`oggenc --downmix --resample 4000 -q -1 file.wav`	6	400	Very low bandwidth
`oggenc --downmix --resample 6000 -q -1 file.wav`	8	300	Mobile phone quality
`oggenc --downmix --resample 8000 -q -1 file.wav`	9	265	Slightly better than mobile phone
`oggenc --downmix --resample 8000 -q 0 file.wav`	13	184	Clean 3 kHz bandwidth sound
`oggenc --downmix --resample 11025 -q -1 file.wav`	14	170	Toy application quality
`oggenc --downmix --resample 11025 -q 0 file.wav`	17	140	Cleaner toy application
`oggenc --downmix --resample 11025 -q 1 file.wav`	20	119	Clean 5 kHz bandwidth sound
`oggenc --downmix --resample 16000 -q 1 file.wav`	24	99	Wideband (7 kHz) sound
`oggenc --downmix --resample 22050 -q 1 file.wav`	30	80	10 kHz bandwidth
`oggenc --downmix --resample 44100 -q 0 file.wav`	36	66	Full bandwidth (15+ kHz)
`oggenc --downmix --resample 44100 -q 1 file.wav`	42	57	Cleaner sound
`oggenc --downmix --resample 44100 -q 2 file.wav`	48	50	Close to transparent
`oggenc --downmix --resample 44100 -q 5 file.wav`	60	40	CD quality (but mono)

The "‑‑downmix" option converts sound to mono and the "‑q" slider works as in Chapter 3.2. Unfortunately, many graphical front‑ends are missing the "‑‑resample" option, which is a very powerful parameter for speech compression. If your program misses this slider, please compress your files from oggenc's command line mode.

3.4. Post‑Processing with VorbisGain

Although the original CD standard defines a reference loudness that should be used for CD recording, current CDs no longer follow that standard. Because of this unequal loudness of CDs, songs compressed from CDs may have wildly differing subjective loudness, a phenomenon that is known to most users of portable digital audio players.

VorbisGain is a program that uses a psycho‑acoustical model to determine an actual, perceived loudness of an audio track or album, and tags this information to the headers of an Ogg Vorbis file. This makes it possible for a VorbisGain aware decoder to adjust its output volume accordingly, and thus equalize loudness differences of songs without a need to recompress them.

Using VorbisGain on all Vorbis songs will enhance the user's listening experience because there is no need to turn volume up and down between songs. Many Ogg Vorbis playback programs and plug‑ins for PC's as well as all VLSI Solution's Ogg Vorbis decoder chips are compatible with VorbisGain.

4. MP3 to Ogg Vorbis conversion by using CDex

CDex is Open Source Digital Audio CD Extractor that has more than 32,000,000 downloads of Windows operating system users. The program is mainly intended for CD ripping but can also be used for transcoding.

CDex has Ogg Vorbis encoder in default installation. However, the default encoder is lacking of settings for speech quality transcoding. Therefore, for speech application it is also installed command line based *oggenc* to be controlled via CDex user interface.

If you want to use CDex for CD quality transcoding please continue reading from Chapter 4.1. But if you intend to use CDex for speech quality transcoding please jump to Chapter 4.2.

4.1. CD audio quality transcoding by using CDex

4.1.1. Install CDex

Download and install CDex binary from http://sourceforge.net/projects/cdexos/
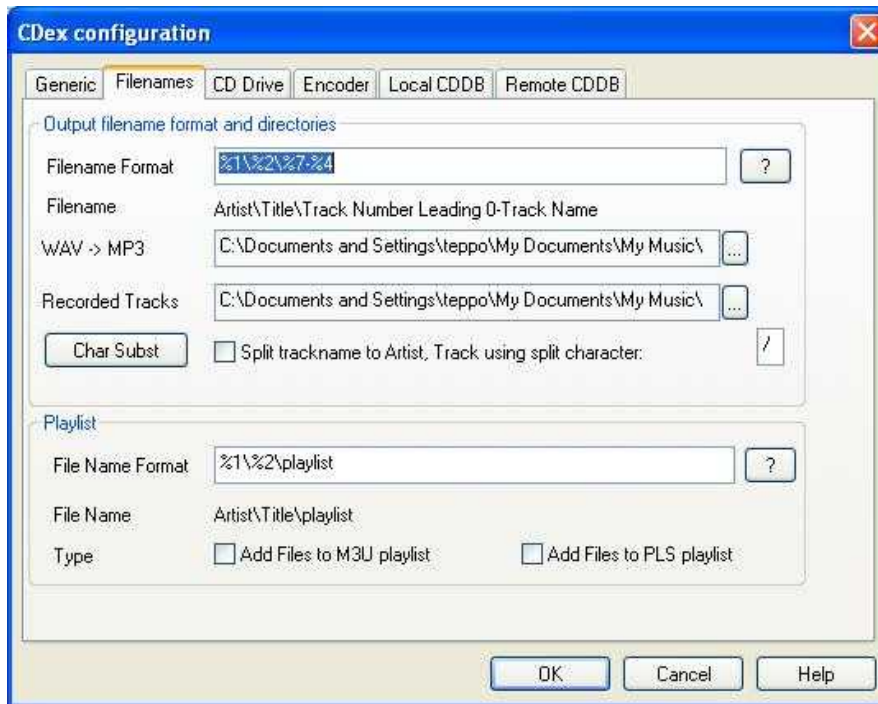
4.1.2. Change CDex settings

Start Cdex

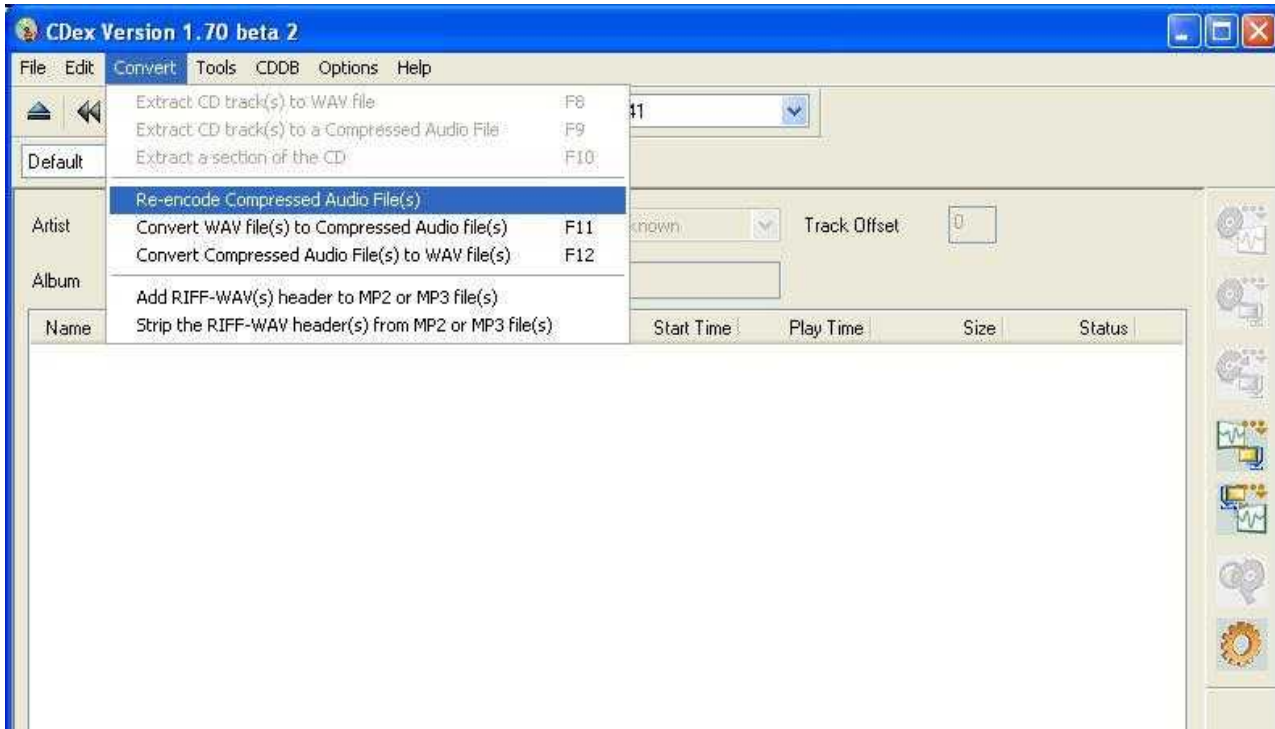Change default converter to Ogg Vorbis under **Options** - > **Settings** (F4) - > **Encoder**
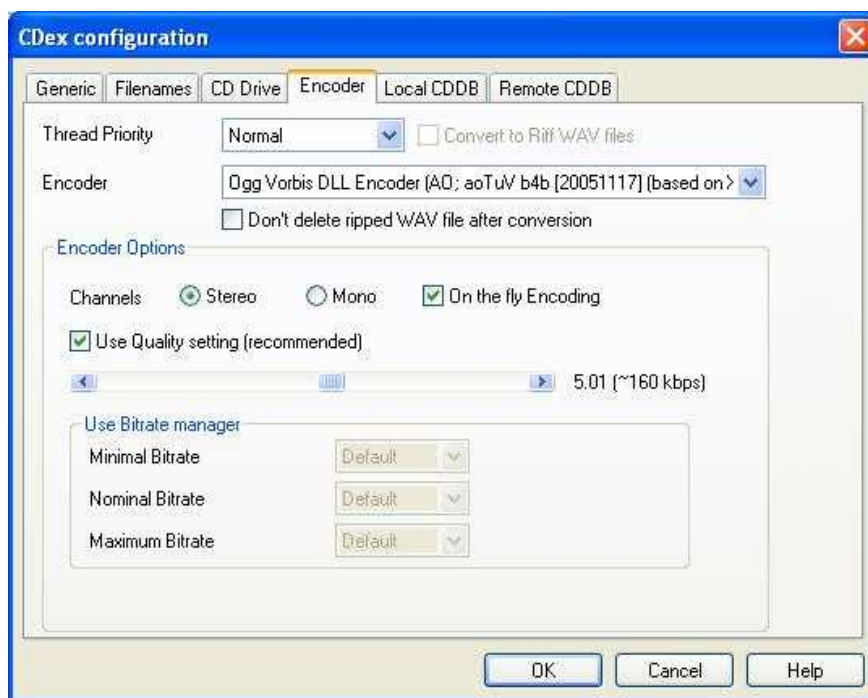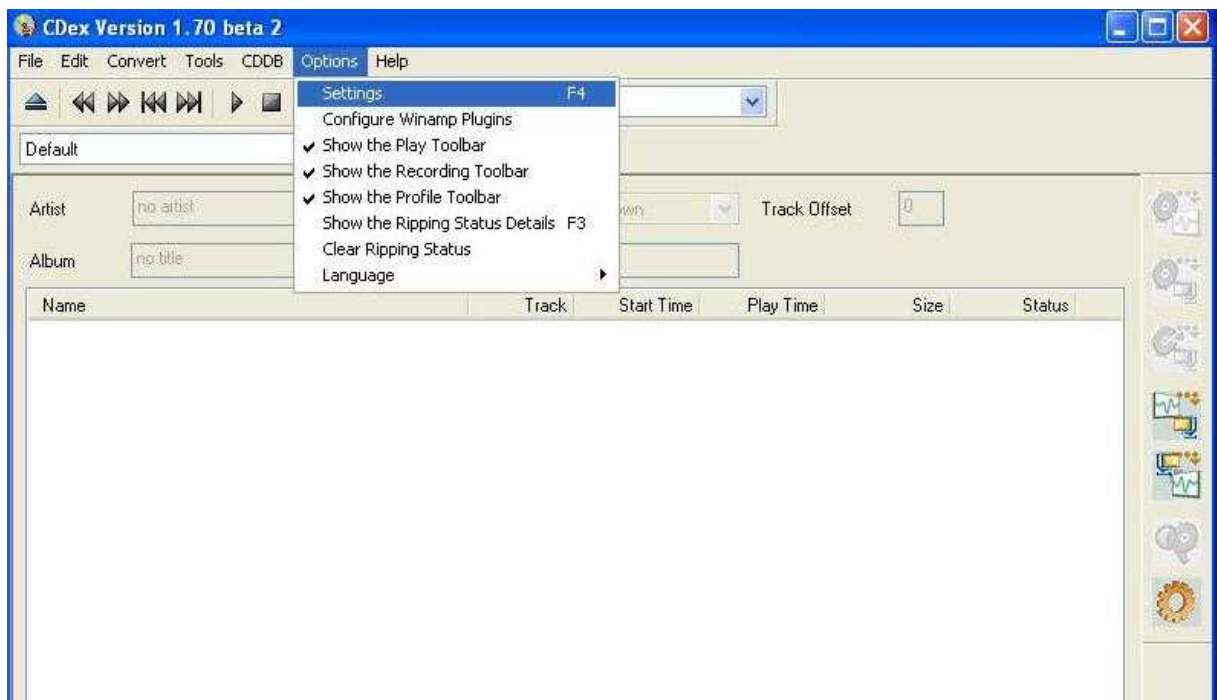
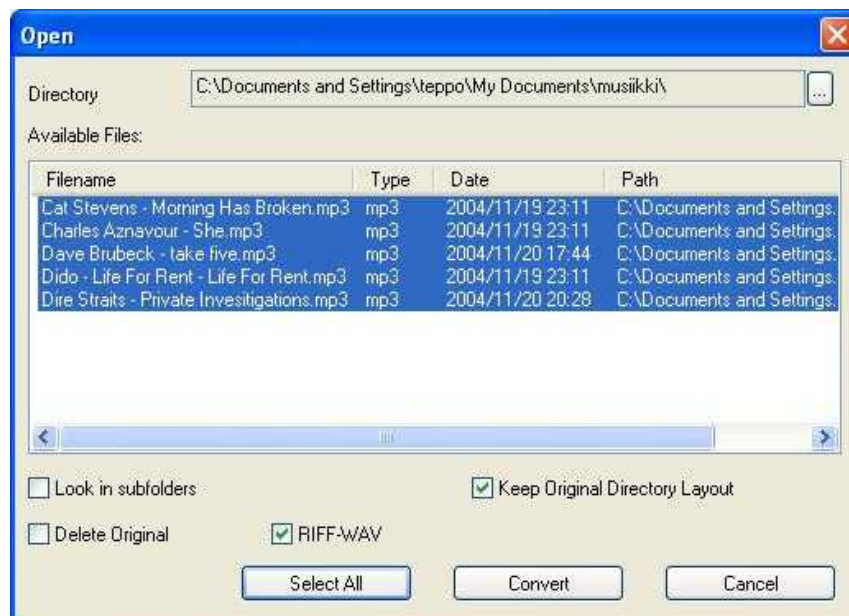Change the target folder of the Ogg Vorbis files
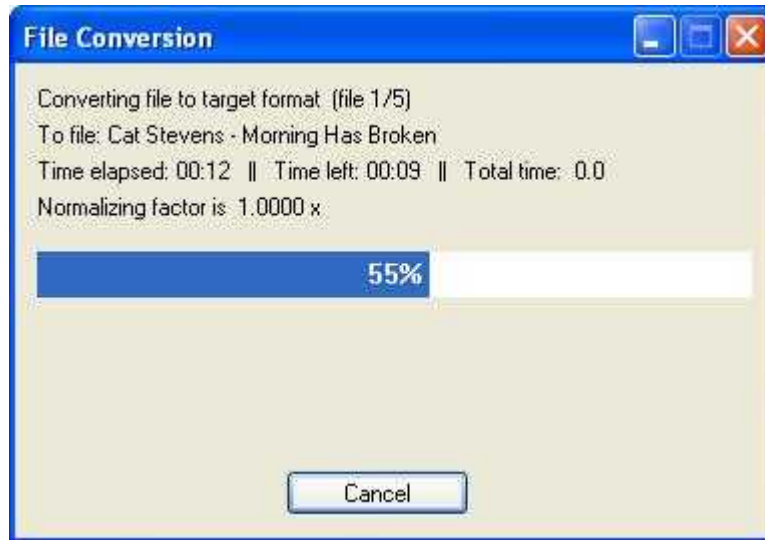
4.1.3. Conversion

Select files to convert under **Convert** - > **Re-encode Compressed Audio File(s)**

Select the files you want to convert and press **Convert**.

4.1.4. Wait conversion result

Files are converted to Ogg Vorbis format (*.ogg).

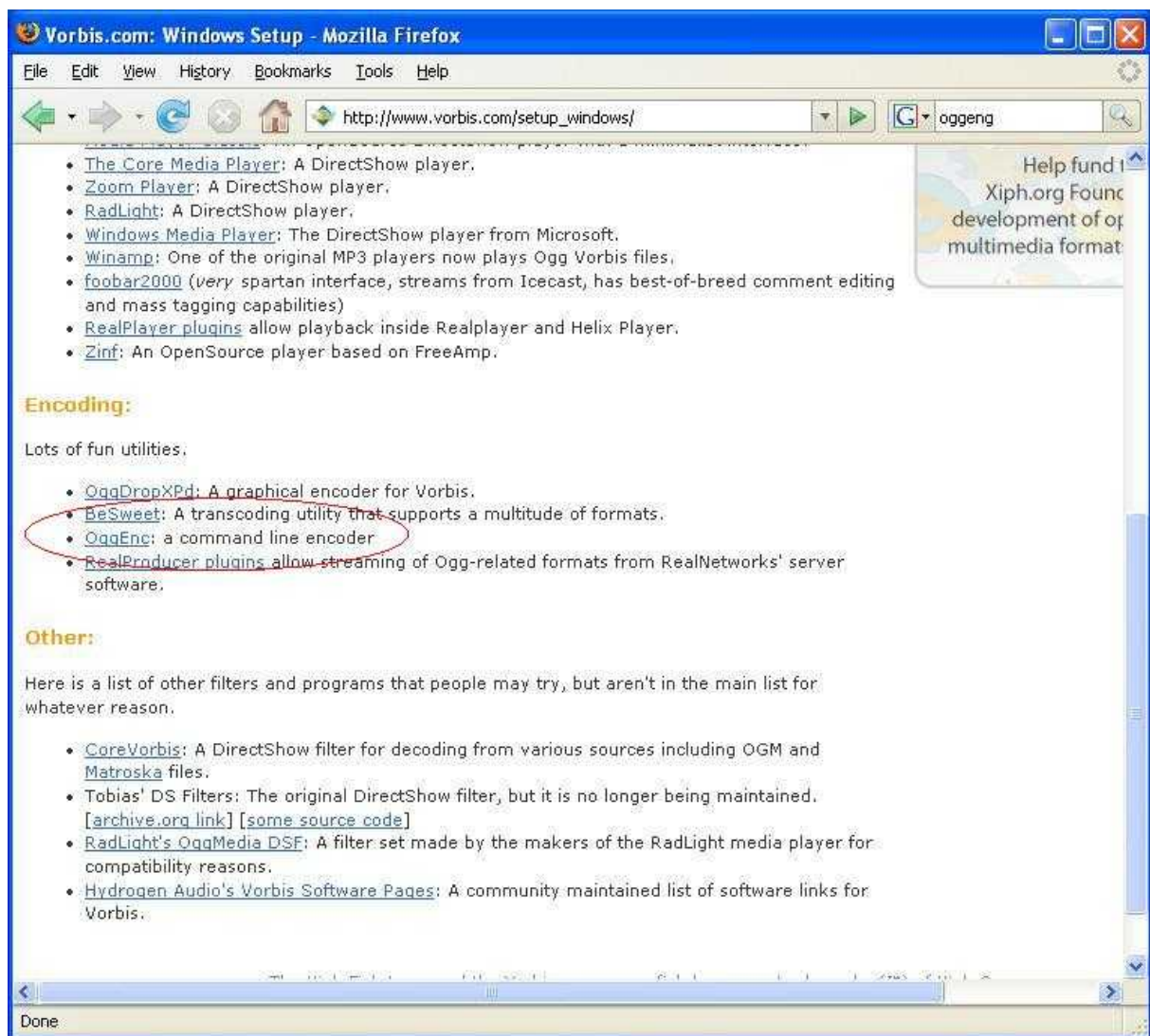Files are locating in the directory you defined in the settings.

## 4.2. Speech quality transcoding by using CDex

4.2.1. Install CDex

Download and install CDex binary from http://sourceforge.net/projects/cdexos/
You do not need to re-install if you have gone through Chapter 4.1.

4.2.2. Install OggEnc

Download binary from http://www.vorbis.com/setup_windows/

You will be redirected to a new site where you can select the encoder binary based on architecture of your computer. Safe selection is Generic.
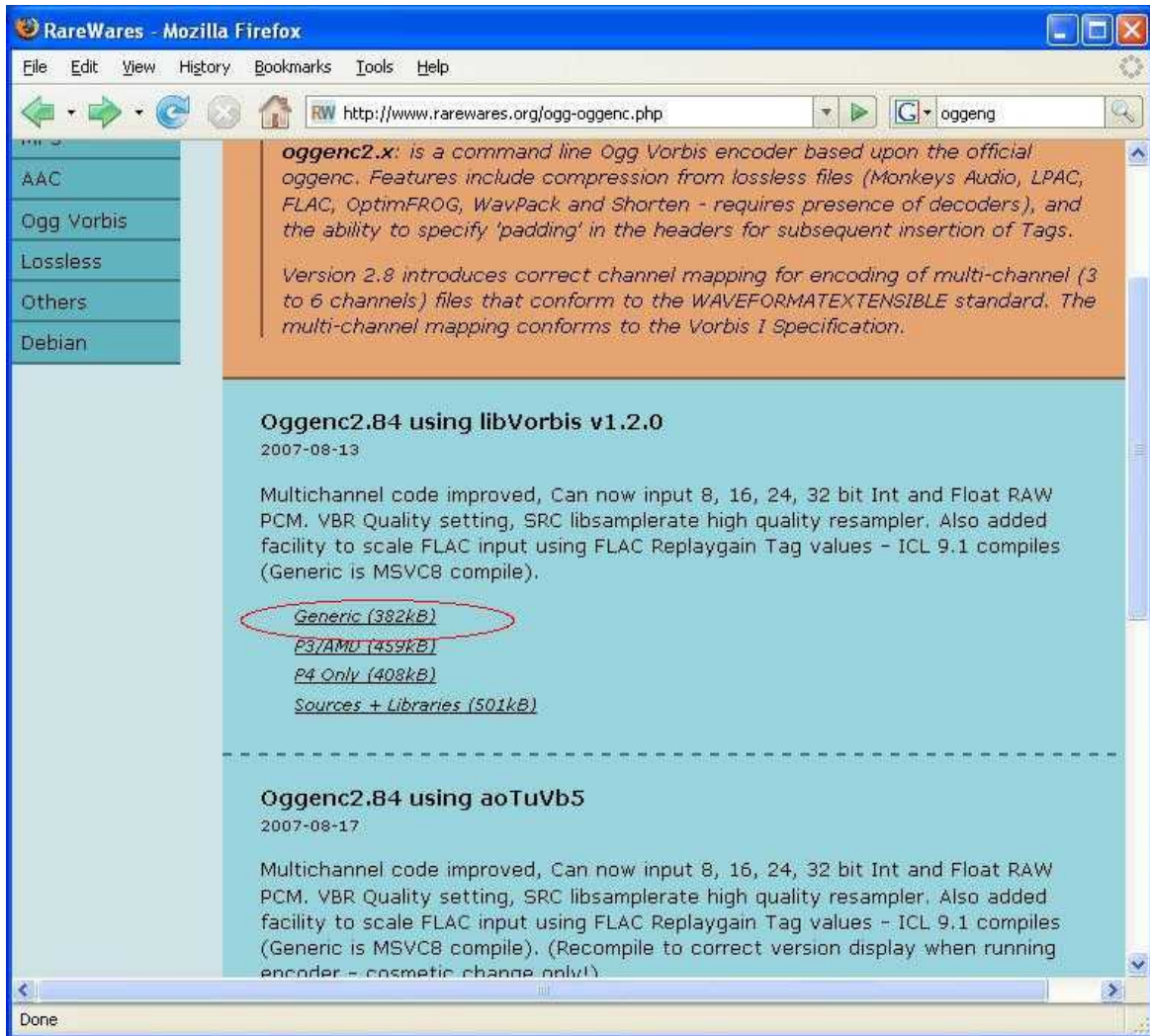
4.2.3. Download

Save the file to disk and extract OggEnc executable from the compressed zip achive.
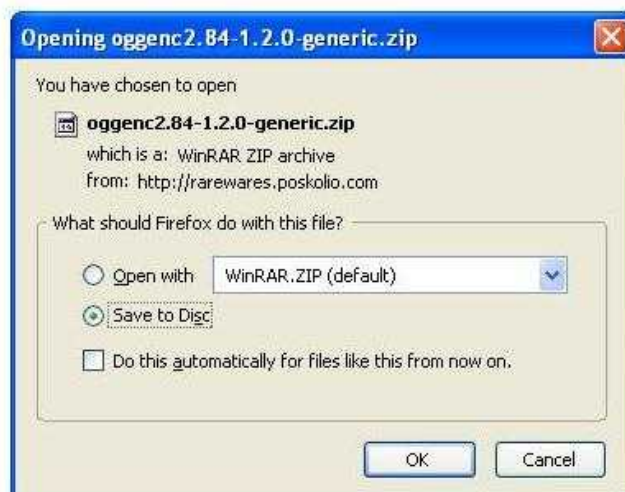
4.2.4. Manually install the binary

Create oggenc directory to your default program files directory and put the oggeng2 binary there.
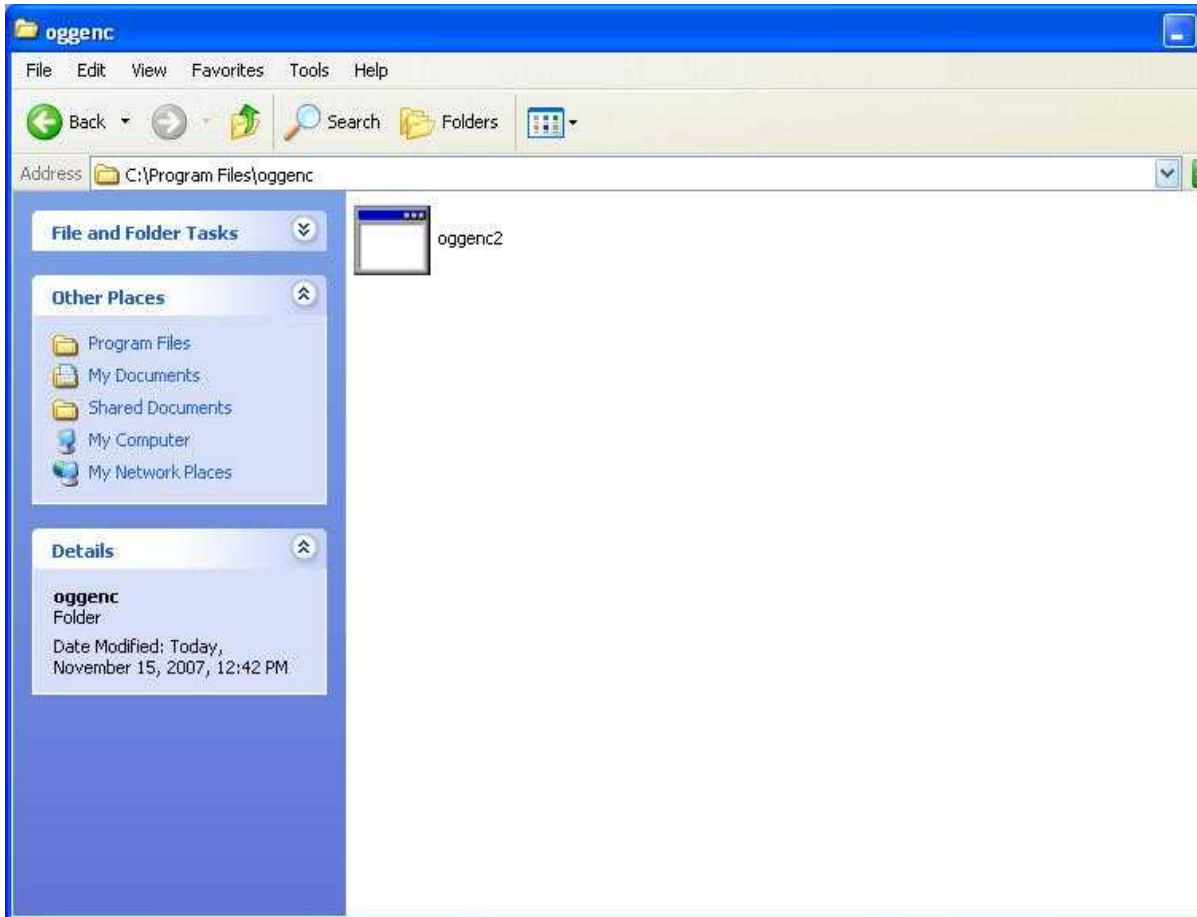
4.2.5. Testing

You can test that oggenc works by opening command prompt (Start - > Programs - > Accessories - > Command Prompt) and writing

oggeng2 - h

to the window

```
C:\Program Files\oggenc>dir
 Volume in drive C has no label.
 Volume Serial Number is BC4F-4C2D

 Directory of C:\Program Files\oggenc

11/15/2007  12:42 PM    <DIR>          .
11/15/2007  12:42 PM    <DIR>          ..
08/13/2007  02:57 PM         1,454,080 oggenc2.exe
               1 File(s)     1,454,080 bytes
               2 Dir(s)   5,784,317,952 bytes free

C:\Program Files\oggenc>oggenc2 -h
OggEnc v2.84 (libvorbis 1.2.0)
(c) 2000-2005 Michael Smith <msmith@xiph.org>
 & portions by John Edwards <john.edwards33@ntlworld.com>

Usage: oggenc2 [options] input.wav [...]

OPTIONS:
 General:
  -Q, --quiet          Produce no output to stderr
  -h, --help           Print this help text
  -r, --raw            Raw mode. Input files are read directly as PCM data
  -F, --raw-format=n   Set format for raw input. Default is 1 for
                       Standard PCM; use 3 for IEEE Float.
  -B, --raw-bits=n     Set bits/sample for raw input. Default is 16,
                       32 assumed for IEEE Float.
  -C, --raw-chan=n     Set number of channels for raw input. Default is 2
  -R, --raw-rate=n     Set samples/sec for raw input. Default is 44100
  --raw-endianness     1 for bigendian, 0 for little (defaults to 0)
  -b, --bitrate        Choose a nominal bitrate to encode at. Attempt
                       to encode at a bitrate averaging this. Takes an
                       argument in kbps. By default, this produces a VBR
                       encoding, equivalent to using -q or --quality.
                       See the --managed option to use a managed bitrate
                       targetting the selected bitrate.
  --managed            Enable the bitrate management engine. This will allow
                       much greater control over the precise bitrate(s) used,
                       but encoding will be much slower. Don't use it unless
                       you have a strong need for detailed control over
                       bitrate, such as for streaming.
  -m, --min-bitrate    Specify a minimum bitrate (in kbps). Useful for
                       encoding for a fixed-size channel.
  -M, --max-bitrate    Specify a maximum bitrate in kbps. Useful for
                       streaming applications.
  -q, --quality        Specify quality between -2 (low) and 10 (high),
                       instead of specifying a particular bitrate.
```
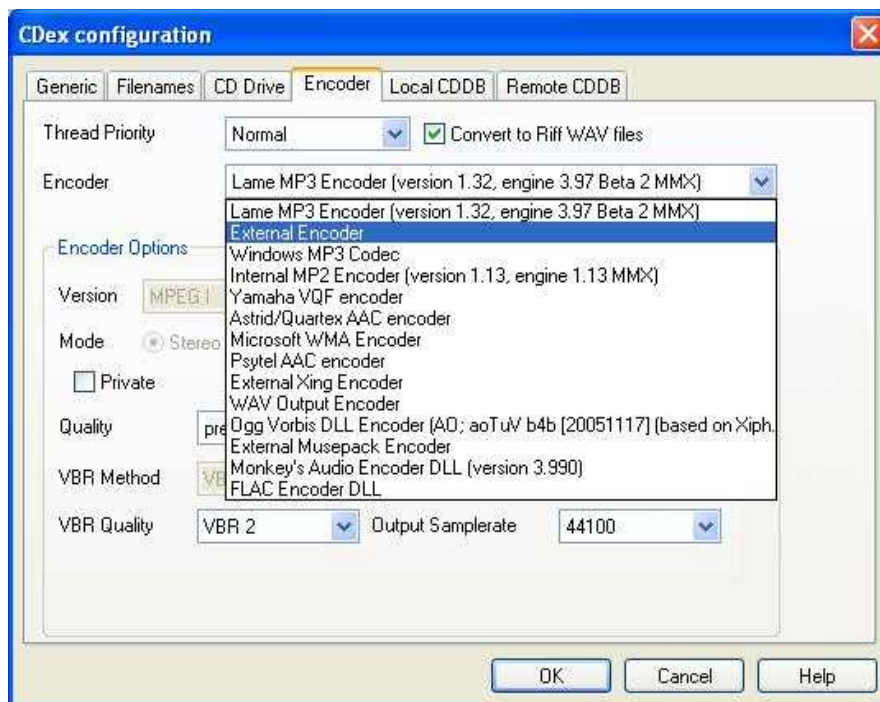
4.2.6. Change CDex settings

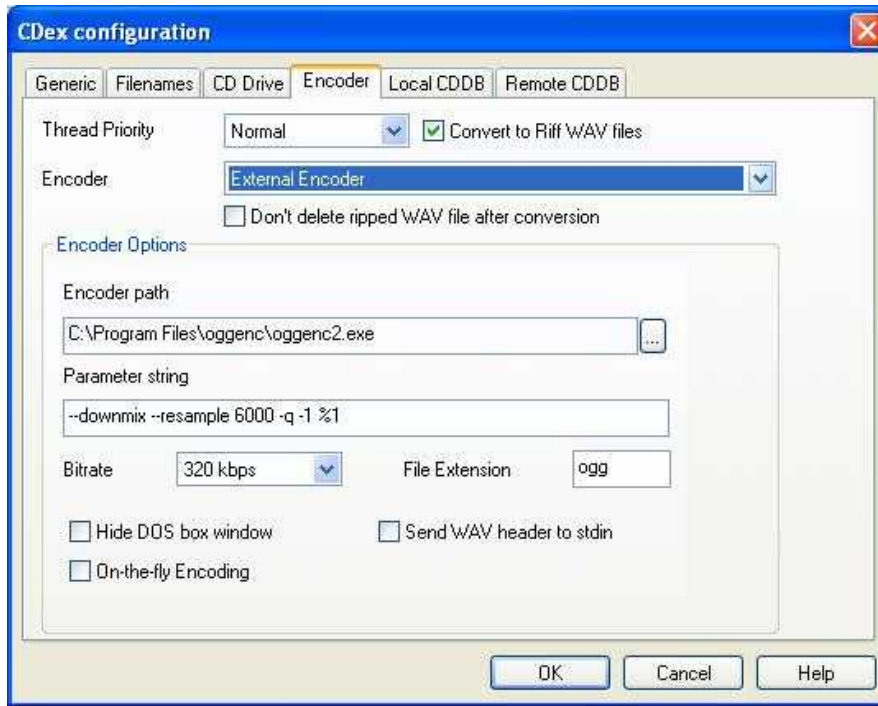Start CDex (Start - > CDex - > CDex) and change the default encoding parameters.
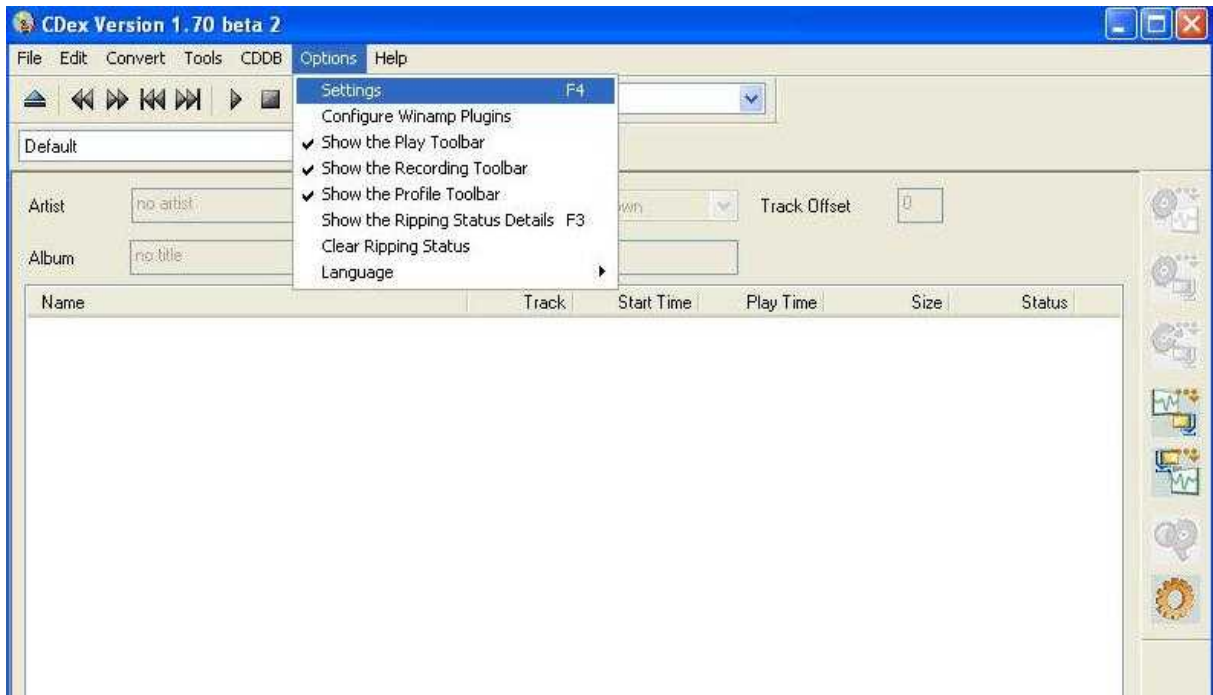
Select Encoder from the sub menu and change it to **External Encoder**.

Set parameters for the external encoder as shown below. Set Parameter string as described in table of Chapter 3.3. For instance following will set 8 kbit/s and downsampling to new sample rate of 6 kHz.

The Encoder path is the one you have put your OggEnc encoder (Chapter 4.24).

Before conversion, please set the default output directories as well. You do not need to do this step if you have gone through Chapter 4.1.
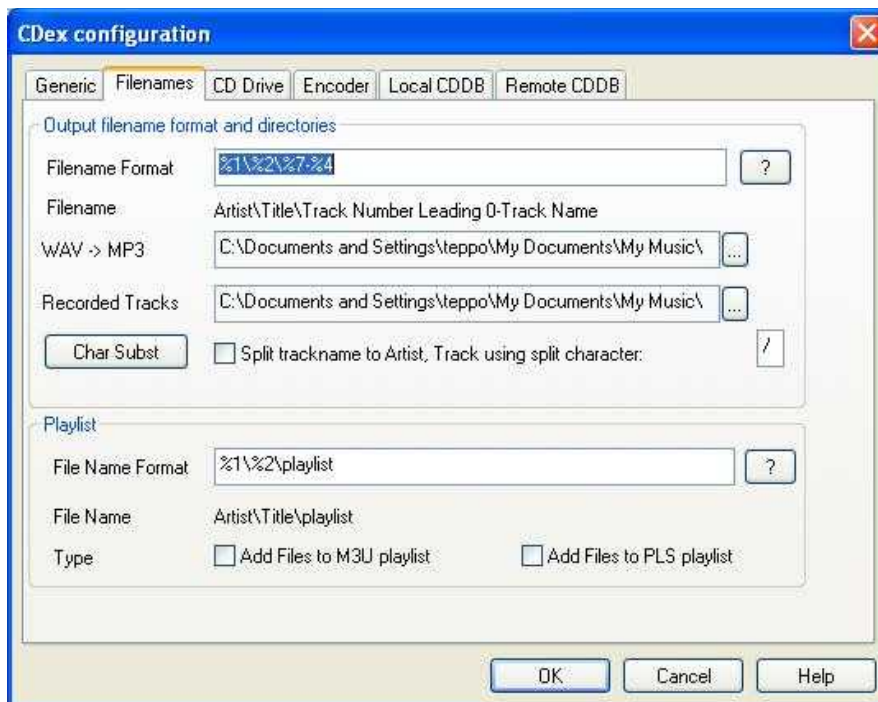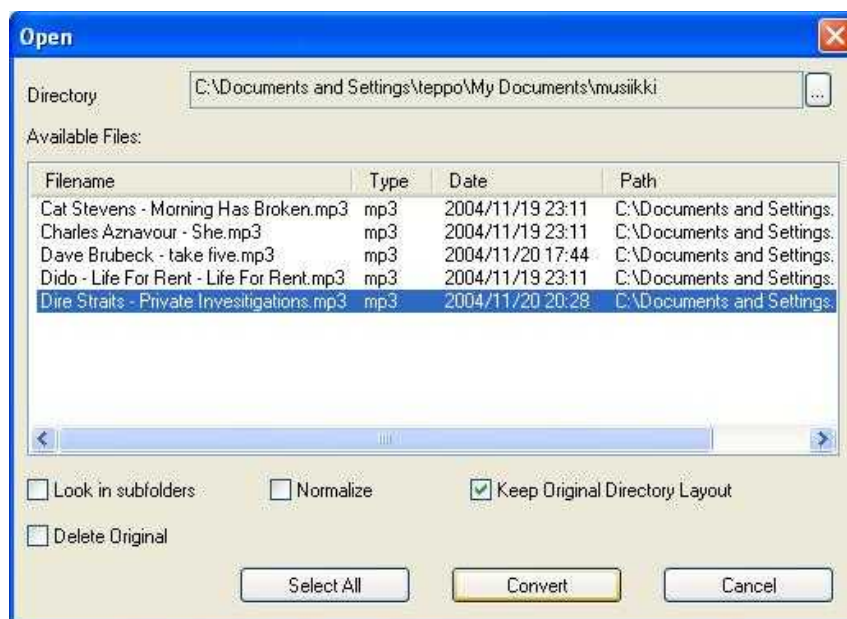
4.2.7. Conversion

Now you are ready to transcode. It is needed to do it in two steps, first from MP3 to WAV, then from WAV to Ogg Vorbis by using external *oggenc* encoder.
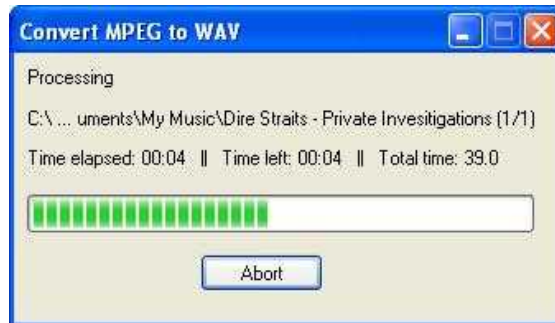
First select files to convert to WAV.
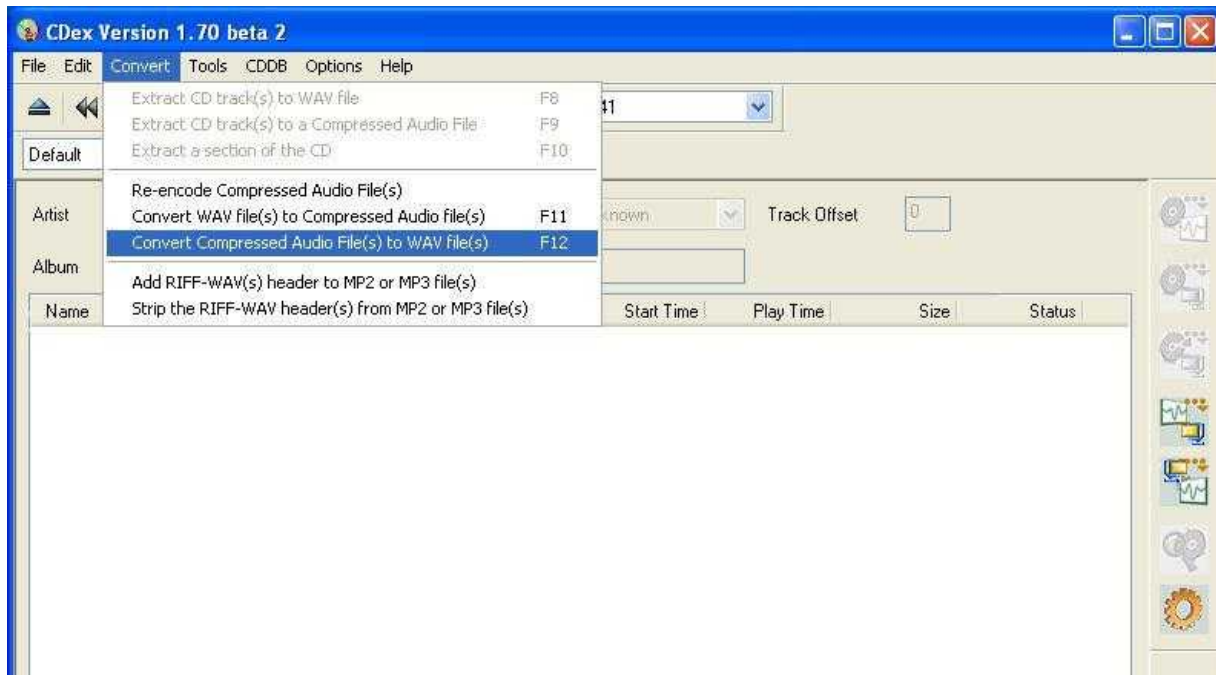
Then select converted files from the directory containing the mp3 files.
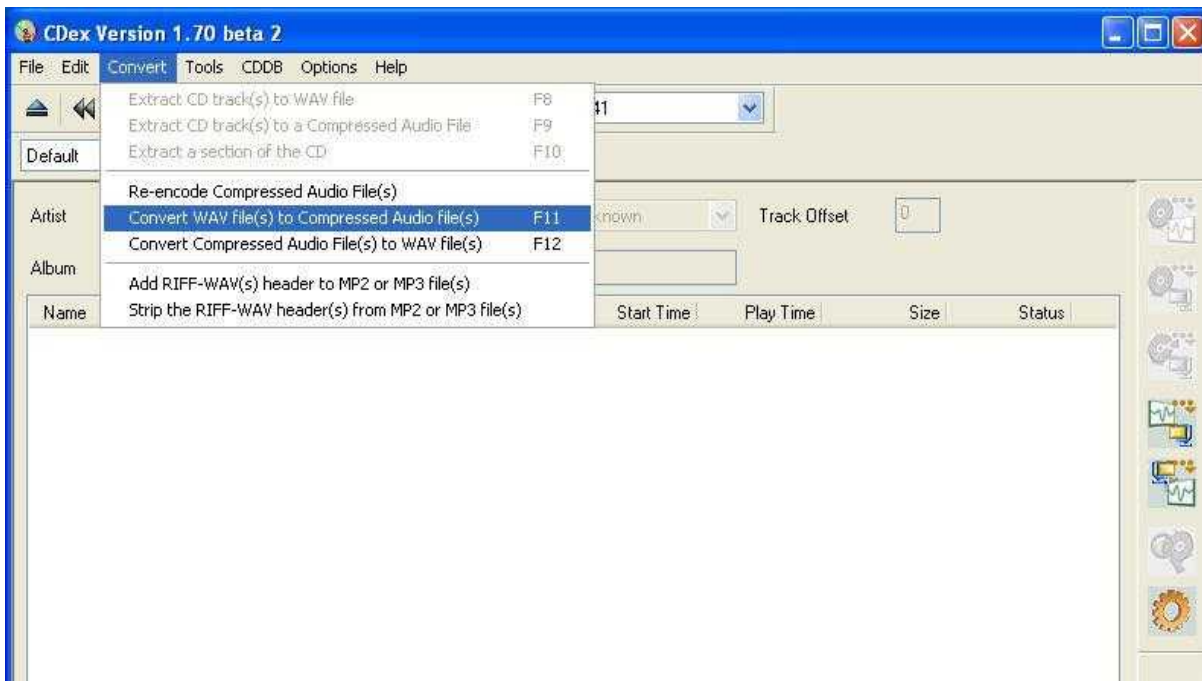

You can select several files if you want. Once ready press **Convert**.

Wait until the conversion is ready. It may take some time if you have selected many files.
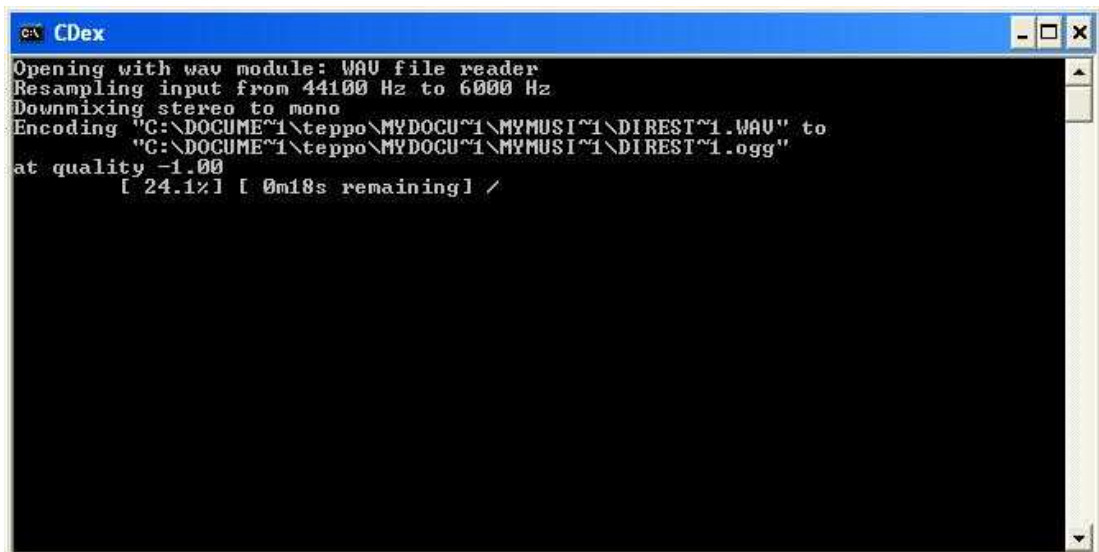

Next you need to convert WAV files to Ogg Vorbis files. At the same time you will decrease the sample rate and bit rate as per your parameter string.

Pick the WAV files to be converted to Ogg Vorbis from the directory that was defined as output directory of the conversions. If you check the Delete Original, the WAV files will be deleted.
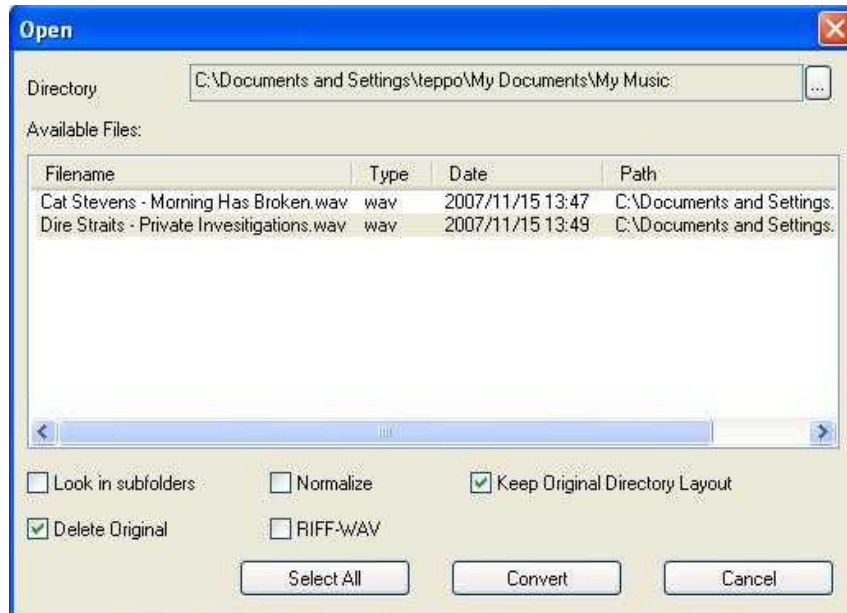
You will see *oggenc* working

You are now ready.

4.2.8. Verify

Start your favourite audio player and verify that bit rate and sample rate are correct.

5. Contact Information

VLSI Solution Oy
Entrance G, 2nd floor
Hermiankatu 8
FIN‑33720 Tampere
FINLAND
Fax: +358‑3‑3140‑8288
Phone: +358‑3‑3140‑8200
Email: sales@vlsi.fi
URL: http://www.vlsi.fi/