Documentation for Github data repository:

The data model
The candidates are sorted by SDSS J name. All of the spectra for each binary candidate are located in the corresponding zipped J######/ directory, where each object has its own directory sorted by the first 6 numbers of its J name. The J######.tar.gz files contain zipped directories for each individual object. The package.tar.gz file contains all of the zipped J######/ directories. That is, you can choose whether to download the data for all the binary candidates, or an individual object of interest.

The filename format encodes the object, date, and observatory. The convention is J######_yyymmdd_obs.ext, where ###### is the abbreviated J name, yyymmdd is the UT time of the observation, obs is the code for the observatory/instrument combination, and .ext is the file format. Some files have a _1_ or _2_ tag in the header. These should be ignored (these were created because when we combined two observations taken within a short time of each other, sometimes the average date had the same value as the date of one observation, so the original file was renamed). The original SDSS spectra have a few different versions: "full" is the full spectrum, no extension is the optical spectrum of H-beta, and "cln" is the result of Eracleous et al. (2012) spectral decomposition. It is best to use the "full" file if you want the SDSS spectrum for something.

The observed spectra:
There is a header associated with each spectrum. It has the following format:

SDSS J153636.22+044127.0
UT 2008/04/07 SDSS
 4806.00 second exposure
3849 points

The first line tells the SDSS J name with RA and DEC formatting. The second line tells the UT date of the observation. The third line gives the exposure length. The fourth line gives the number of points in the spectrum. After the header, each observed spectrum has two columns:

Wavelength
Flux density

The wavelengths are given in A in the rest frame and all flux density arrays are in units of 10^-17 erg s^-1 cm^-2 A^-1. Many of the files are in the 2ca file format, which is simply an ASCII two-column format. The filename conventions are J######_YYYYMMDD_OBS.2ca, where J###### is the SDSS J name convention with numbers represent the first six digits of the object RA, the YYYYMMDD tag gives the UT date of the observation, and the OBS tag identifies the observatory. The telescope and instrument configurations for these OBS identifiers can be found in Eracleous et al., 2012, ApJS, 201, 23 and Runnoe et al., 2015, ApJS, 221, 7 and have the following values:

SDSS
HET
MDM
KPNO
APOb
APOr
PalB
PalR

There are three SDSS spectra per object. The J######_YYYYMMDD_OBS_full.2ca file has the full SDSS spectrum. The J######_YYYYMMDD_OBS.2ca file has a concatenated spectrum with extreme wavelengths trimmed. The J######_YYYYMMDD_SDSS_cln.2ca spectrum includes the spectrum for the broad H-beta only, after the contaminating emission-line and continuum components have been removed. This spectral decomposition was done by Eracleous et al. (2012) and does not match that of Runnoe et al. (2015), with the differences described in the appendix of the latter work.

The Runnoe et al. (2015) spectral decompositions:
The spectral decomposition methodology is described in Runnoe et al., 2015, ApJS, 221, 7. The final data product is an ASCII table where columns represent different components of the fit. Filename conventions match those for the observed spectra, with the _mod.txt ending. The columns are:

Wavelength
Flux density
Total spectral model
Power law
Fe II
He II λ4686
Broad H-beta
Narrow H-beta
[O III] λ4959
[O III] λ5007

The wavelengths are given in A in the rest frame and all flux density arrays are in units of 10^-17 erg s^-1 cm^-2 A^-1. All of the flux arrays are sampled at the given wavelength array. The wavelength and flux arrays match those given in the associated two-column observed spectra (i.e. 2ca format).

As an example, you can get the broad H-beta only spectrum by taking the following:

Flux density - Total model + Broad Hβ

You could plot a spectrum (in IDL) like as follows. This example will show the observed SDSS spectrum, the spectral decomposition result in red, the data for the broad Hβ, and the best-fitting broad Hβ model in red.

IDL> $pwd
J153636/
IDL>
readcol,'J153636_20080407_SDSS_full_mod.txt',wave,flux,model,pl,fe2,host,bhb,nhb,o3s,o3l,/silent
IDL> plot,wave,flux,psym=10,xtit='Rest Wavelength ['+cgsymbol('Angstrom')+']',ytit='!8f!X!d'+cgsymbol('lambda')+'!n [10!u-17!n erg s!u-1!n cm!u-2!n '+cgsymbol('Angstrom')+'!u-1!n]',charsize=2
IDL> oplot,wave,model,color=cgcolor('red')
IDL> oplot,wave,flux-model+bhb
IDL> oplot,wave,bhb,color=cgcolor('red')