

patentlin

version 2.0

USER MANUAL

Edition: August 3, 1998

TABLE OF CONTENTS

1. Introduction	1
1.1 PatentIn v. 2.0	1
2. System Requirements and Installation	2
2.1 System Requirements	2
2.2 Pre-Installation Information	2
2.2.1 Definitions	2
2.2.2 Data Entry Format.....	3
2.3 Installation	4
2.3.1 Internet Download – INSTALL.EXE (www.uspto.gov).....	4
2.3.2 Floppy Diskette-SETUP.EXE.....	4
2.3.3 Installation Procedures	4
2.4 INDEXING DATABASE FILES	5
2.5 UNINSTALL	5
2.6 Start the Program	5
3. Structure Organization	6
3.1.1 Application Data Entry Screens	6
3.1.2 Sequence Functions Data Entry Screens.....	7
3.2 Maneuvering in a Data Entry Screen.....	7
3.2.1 Data Entry Screen Cursor Control Keys	7
3.2.2 Cursor Control Keys for Lists	7
3.3 PatentIn v. 2.0 Limitations	8
4. Getting Started - Choose An Application	9
4.1 Create.....	9
4.2 Open.....	9
4.3 Save As.....	9
4.4 Delete.....	10
4.5 Exit	10
5. The Main Menu and Quick Guide	11
5.1 Using the Main Menu	11
5.1.1 Applicant Data	11
5.1.2 Sequence Functions.....	11
5.1.3 Exit.....	12
5.1.4 Help.....	12
5.2 The Quick Guide.....	12
5.2.1 OK.....	12
5.2.2 Cancel.....	12
5.2.3 Exit.....	12
5.2.4 Help.....	12
6. Tips for Completing Data Entry Screens.....	14

6.1	Completing Fields in Data Entry Screens	14
6.2	Mandatory Fields in Data Entry Screens.....	14
6.3	Lists With Restricted Vocabulary	14
6.4	Save and Exit Button Behavior Throughout PatentIn Version 2.0	15
7.	<i>Getting Help.....</i>	<i>16</i>
7.1	Levels of Help Available	16
7.1.1	General Help.....	16
7.1.2	Field Specific Help.....	16
7.1.3	Data Entry Screen Help	16
8.	<i>Completing the Application Data.....</i>	<i>17</i>
8.1	Applicant Information.....	17
8.1.1	Applicant Information Data Entry Screen.....	17
8.1.2	Applicant Information Browse Screen.....	18
8.2	Current Application Information.....	19
8.2.1	Current Application Data Entry Screen	19
8.3	Prior Application Information	20
8.3.1	Prior Application Information Data Entry Screen	20
8.3.2	Prior Application Information Browse Screen	22
9.	<i>Completing the Sequence Functions.....</i>	<i>23</i>
9.1	Sequence Data Browse Screen.....	23
9.1.1	Organism	24
9.1.2	Feature	25
9.1.3	Publications	25
9.1.4	Author Information Browse Screen.....	29
9.1.5	Author Information Data Entry Screen.....	30
9.1.6	Library of Previously Defined Publications Browse Screen	30
9.1.7	Import.....	32
9.1.8	Modify	33
9.1.9	Add.....	33
9.1.10	Delete	33
9.1.11	Restore.....	33
10.	<i>Using the Sequence Editor.....</i>	<i>34</i>
10.1	Sequence Editor Screen	34
10.2	Accessing the Sequence Editor	35
10.3	Commands Available from the Menu Bar	35
10.3.1	FILE Menu	35
10.3.2	Attributes	37
10.3.3	SEARCH Menu.....	37
10.4	Exiting.....	38
10.5	Reordering Sequences	39
10.6	Code '000'	39
10.6.1	Purpose	39
10.6.2	How to insert the "000" code.....	40

11. Completing the Feature Information	41
11.1 Feature Information Data Entry Screen.....	41
11.1.1 Data Entry Block.....	42
11.1.2 Other Information Window.....	43
11.1.3 Feature Location Window.....	43
11.2 Choosing a Feature Name/Key	43
11.2.1 Specific Feature Keys.....	43
11.2.2 Additional Information Button.....	44
11.3 Specifying the Feature's Location	44
11.4 Saving the Feature	44
11.5 Receiving an Error or Warning Message When Saving - The Next Action to Take.....	45
11.6 Editing a Feature	45
11.7 Deleting a Feature.....	45
11.8 Exiting the Feature Information Data Entry Screen.....	45
11.9 Annotating a Protein Coding Region	45
11.10 Annotating an Internal Stop Codon in a Protein Coding Region	47
11.11 Example of Sequence Annotation	47
12. Generate an Application Sequence Listing	49
12.1 Generating the Sequence Listing.....	49
12.2 Browse Sequence Listing.....	49
12.3 Printing the Sequence Listing.....	50
12.4 Editing the Sequence Listing	50
12.5 About Sequence Listing Output	51
12.5.1 Numbering Sequence Listing Residues in the Output	51
12.6 Information on How Data is Stored	51
12.7 Final Steps.....	52
13. Reference	53
13.1 Feature Keys - Nucleic Acid	53
13.2 Feature Keys - Amino Acid	56
13.3 Feature Keys - Additional Descriptions	58
13.4 List of Numeric Heading Definitions.....	59

TABLE OF FIGURES

<i>Figure 3-1: PatentIn v. 2.0 Program Organization</i>	<i>6</i>
<i>Figure 4-1: Choose An Application</i>	<i>10</i>
<i>Figure 5-1: PatentIn v. 2.0 Main Menu/Quick Guide</i>	<i>13</i>
<i>Figure 8-1: Applicant Information Data Entry Screen.....</i>	<i>18</i>
<i>Figure 8-2: Applicant Information Browse Screen</i>	<i>19</i>
<i>Figure 8-3: Current Application Data Entry Screen.....</i>	<i>19</i>
<i>Figure 8-4: Prior Application Data Entry Screen.....</i>	<i>21</i>
<i>Figure 8-5: Prior Application Information Browse Screen</i>	<i>22</i>
<i>Figure 9-1: Sequence Data Browse Screen</i>	<i>24</i>
<i>Figure 9-2: Organism Button Options.....</i>	<i>24</i>
<i>Figure 9-3: Library of Previously Defined Publication</i>	<i>26</i>
<i>Figure 9-4: Bibliographic Information Data Entry Screen</i>	<i>27</i>
<i>Figure 9-5: Bibliographic Information, continued Data Entry Screen.....</i>	<i>29</i>
<i>Figure 9-6: Author Information Browse Screen</i>	<i>30</i>
<i>Figure 9-7: Import File Manager</i>	<i>32</i>
<i>Figure 10-1: Sequence Editor.....</i>	<i>34</i>
<i>Figure 10-2: Example of 000 Place Holder.....</i>	<i>40</i>
<i>Figure 11-1: Feature Information Data Entry Screen.....</i>	<i>42</i>

1. Introduction

1.1 PATENTIN V. 2.0

PatentIn v. 2.0 is a network compatible program designed to expedite the process of preparing applications containing nucleic acid and polypeptide sequences. PatentIn v. 2.0 prepares Sequence Listings in adherence to WIPO ST.25.

PatentIn v. 2.0 solicits all the information needed for a sequence application, including information about the applicants, prior patent publications, journal references pertaining to the sequence, and biological information about the sequence and its source. The program is accessible in any of four languages: English, German, Spanish, or French. Wherever possible, PatentIn v. 2.0 allows you to make selections from pick lists, thus standardizing terminology and reducing typing errors. You may enter any part of the data at any time and in any order you wish, and you can easily add, remove, and revise it. You can also save a partially completed application and finish it during a later session.

A Sequence Editor is included as part of PatentIn v. 2.0, allowing you to enter and modify both nucleic acid and protein sequences. Use the Sequence Editor to enter multiple sequences for each application. You can import sequence files created by another editor or word processor (provided they have been stored as ASCII text files) into the PatentIn v. 2.0 Sequence Editor and incorporate them into your application.

After PatentIn v. 2.0 has collected the data necessary for your application, it will assist you in creating an Application Sequence Listing. This is a computer-readable file containing a Sequence Listing prepared in accordance with WIPO ST.25. You can print a complete copy of this file and submit it as the paper version of the application.

2. System Requirements and Installation

2.1 SYSTEM REQUIREMENTS

PatentIn v. 2.0 is available for IBM compatible computers using a Windows™ operating system. The minimum requirements are as follows:

- The minimum system hardware requirement is a 386 processor (or equivalent) and 4 MB of Random Access Memory (RAM). However, for optimum performance, it is recommended that systems have a 486/33mhz processor with not less than 8 MB of RAM.
- The operating system must be Windows 3.1, Windows for Workgroups 3.11, Windows 95, or Windows NT 3.51 or higher.
- Operating systems must have a hard disk drive with 20 megabytes free disk space plus enough file storage space to accommodate all of the sequence files to be imported to the system.

2.2 PRE-INSTALLATION INFORMATION

2.2.1 Definitions

Before you begin the installation, you should know some key words and definitions:

Application: This term is used throughout the program and user manual to denote a discrete Sequence Listing project.

Cursor: The marker that indicates where you are on the computer terminal screen.

Button: In a graphics display, a window control that performs a function when activated with a mouse click or keystroke. (e.g., OK, Cancel, Exit are typical buttons.)

Browse (Browser): A list of items that can be scrolled from top to bottom with keyboard controls or a mouse. In a graphics display, a browser is a separate control. (See Listbox)

Child Window: In a graphics display, a window within a window.

Clipboard: In a Windows™ operating system, an area in computer memory where data is stored when it either “cut” or “copied.” The contents of the clipboard can be “pasted” (copied to) anywhere in the operating system. (i.e., a name can be “cut” from the notepad editor and “pasted” into the name field of PatentIn.). Because all sequences in PatentIn MUST BE SCREENED, the clipboard CANNOT be used to paste sequences into the “Sequence Editor.”

Combination Box (or Combobox): In a graphics display, a control that is a combination data entry field and list. These controls are characterized by a down-arrow to the right of the window. When this arrow is activated with the mouse, the list drops down. This list behaves like a Listbox (see Listbox, this paragraph.). Once a selection is made, the list disappears and the selection appears in the data entry

window. In PatentIn Version 2.0, all Comboboxes are “read-only.” They must be filled from the associated lists.

Control: In a graphics display, an item that is displayed to perform a specific function. Buttons perform actions when activated, text boxes display text, edit boxes accept data entries, listboxes displays lists of items to browse.

Data Entry Box (or Window): In a graphics display, a control in which data can be entered by typing or “pasting” from the Windows™ clipboard.

DOS: The Disk Operating System.

Dialog Box: In a graphics display, a child window within a main window designed to contain controls (e.g., see buttons, text boxes, radio menus, window, child window.)

Directory: A table of contents for a disk. Separate directories can be created on your disk drive to store files in an organized manner.

Listbox: In a graphic display, a browser control. (See Browse)

Menu: A list of actions to be executed when selected and activated.

Root Directory: The highest directory level on your hard drive (usually designated as C:\ or D:\) or an external file server on a network.

Subdirectory: A directory within a directory.

Default: The option taken if you do not specify a value or make a decision. The default directory is the directory in which the programs (including the operating system) look for files, unless you specify another directory.

Path: A combination of the disk drive designation and the directory name that tells the computer where to look for a file if it is not contained in the current directory.

Radio Menu: In a graphics display, a menu of options that are connected to operate like the station buttons on a typical radio. In this application, radio menu selections, once set, are activated with an **OK** button, which is always the default action if you press **Enter**.

Text Box: In a graphics display, a control to display text.

Windows™: A reference to the Microsoft graphical user interface (GUI) operating system.

Window: A rectangle on a video screen to display data. These include the main window, child window (a window within a window), or a data entry or display box in a Windows™ based computer application.

2.2.2 Data Entry Format

When you type installation commands, you may use upper- or lower-case letters. If your hard disk drive is not drive C, use the letter of its designation in place of C for the installation instructions in **Section 2.3** “Installation.”

PatentIn v. 2.0 installs the PatentIn v. 2.0 files into a subdirectory created with that name. Sequence files and the associated Sequence Listing file are stored in subdirectories of the

installation directory created with the application filename for each new application project. Since meaningful and systematic names help organize the storage of the application files on your computer, uniform conventions for file names are recommended.

PatentIn v. 2.0 stores sequence data in files with the extension .GBS. PatentIn v. 2.0 also assigns the computer-readable file (the Sequence Listing) an .APP extension. Since these two file extensions have specific meanings for PatentIn v. 2.0, they should not be assigned to any files in your application directory other than application and sequence files created by PatentIn v. 2.0.

2.3 INSTALLATION

2.3.1 Internet Download – INSTALL.EXE (www.uspto.gov)

INSTALL.EXE is the PatentIn Version 2.0 single file installation file that can be downloaded from the Internet. Store this file in any directory on your PC or a file server. The installation may proceed from this directory. The INSTALL.EXE file does not need to be in the destination directory.

2.3.2 Floppy Diskette-SETUP.EXE

SETUP.EXE is the PatentIn Version 2.0 two-part file available on two 3½ inch floppy diskettes. Insert disk one into a local floppy drive to begin installation. The user will be prompted to insert the second disk midway through the installation.

2.3.3 Installation Procedures

For Windows 3.1, activate the FILE option from the main window menu bar, then select RUN. Type in the drive and full path name of the installation file (INSTALL.EXE or SETUP.EXE), or use the browse button to select the file. Activate OK after the file name is set in the window. Follow the screen directions. PatentIn creates a group window and installs the corresponding system icon (BOOKS) in that group window.

For Windows 95 and Windows NT, use the START button, then select RUN, then follow installation instructions described above for Windows 3.1. The PatentIn Version 2.0 icon (BOOKS) is installed in the Task Bar under Programs.

NOTE: Mandatory for Windows NT users only: In order for the program to run properly, the PatentIn properties **must** be adjusted to run in separate memory space. After the Windows NT installation, set the properties of the PatentIn Version 2.0 file to “Run In Separate Memory Space” by following these directions:

1. Activate the START button, select “Settings,” then select the option “Taskbar.”
2. In the Task Bar, select the “Start Menu Properties” folder, then click on the “Advanced” button.

3. Double click on the “Programs” icon, which will display all of the program applications under your START menu. Navigate to the PatentIn Version 2.0 file, “**right** click” the mouse on the file, then select “Properties.”
4. Click into the second tab titled “Shortcut.” Click in the box that says “Run in Separate Memory Space.” Exit by hitting the “OK” button.

2.4 INDEXING DATABASE FILES

PatentIn Version 2.0 indexes all database files during the installation. If later, the system experiences an abnormal termination while PatentIn Version 2.0 is running (e.g., a power failure), the system’s databases may need to be reindexed in order for data to be displayed and printed accurately.

All users must exit PatentIn before indexing can proceed. After all users have exited the program, use the File Manager (in Windows 3.1) or Explorer (in Windows 95 or Windows NT) to activate the file W_INDEXE.EXE. (For convenience, the corresponding DOS utility, INDEX.EXE is present as well.) Indexing requires only a few minutes.

2.5 UNINSTALL

“Uninstall” utilities are installed in the PatentIn Directory. They consist of two files that must be run in sequence in order to remove all subsequently created subdirectories, if application projects have been added to PatentIn. Use the File Manager (in Windows 3.1) or Explorer in Windows 95 and Windows NT to activate the file REMOVPRJ.EXE. This file removes all subdirectories and most of the installation. Then execute UNWISE.EXE. This will remove the icons from the Windows 95 and Windows NT installations and the balance of the files. In Windows 3.1, if the icon and/or group are still on the desktop, delete them by highlighting and pressing the DELETE key.

2.6 START THE PROGRAM

5. Execute the program by doing one of the following.
 - In Windows 3.1, double click on the PatentIn icon.
 - In Windows 95 or NT, click the START button, select the Programs directory, click on the PatentIn Version 2.0 icon.
6. The program will start and prompt you to select a language. Select the desired language and click “OK.”
7. Select the “Create” button to enter an application project’s Filename and File Reference. After “**Save/Exit**,” select the desired project and click “OK” button.
8. Follow steps 1 through 3 on the Quick Guide to create a sequence listing.

3. Structure Organization

PatentIn v. 2.0 contains a series of data entry screens that request annotated and biological information for the sequence in your patent application. The Application Data entry screens collect data on the applicant and the patent application, while Sequence Functions data entry screens collect data on the nucleic acid or protein sequence. Access to the Application Data and Sequence Functions data entry screens is provided through both the *Main Menu* and the menu bar.

Figure 3-1 is a schematic representation of the PatentIn v. 2.0 structure. This diagram illustrates the relationships among the *Main Menu* items, the data entry screens, and the related sub-screens.

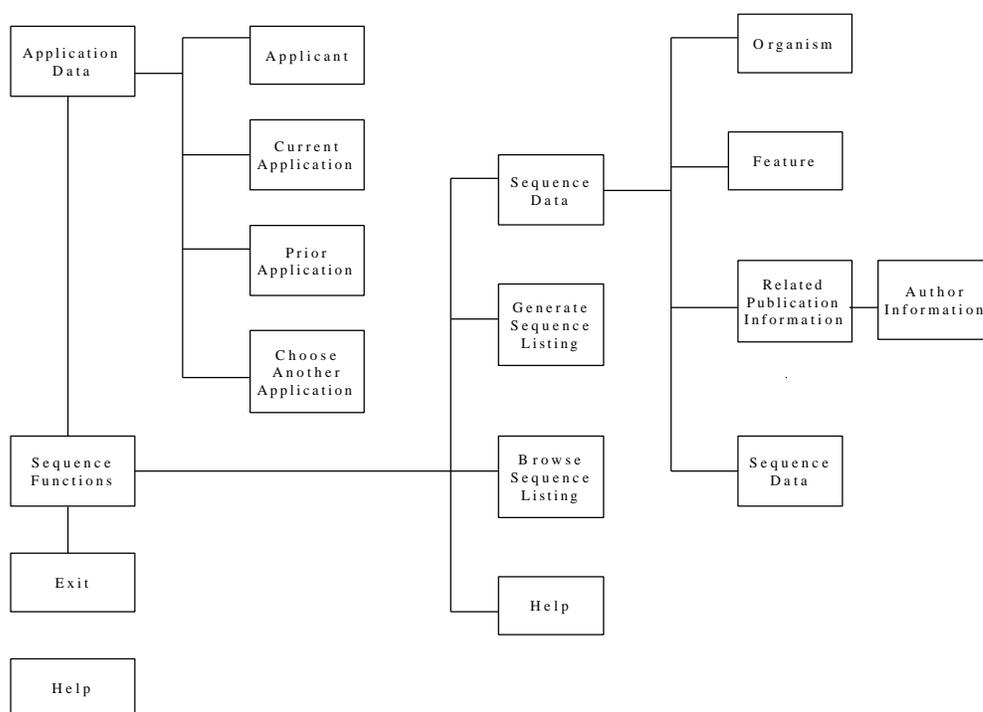


Figure 3-1: PatentIn v. 2.0 Program Organization

3.1.1 Application Data Entry Screens

<i>Data Entry Screen Name</i>	<i>Information Recorded</i>
Applicant	Applicant's name, address, and related information.
Current Application Information	Current application information: Title, Application Number, Filing Date, File Reference
Prior Application Information	Prior patent applications: Application number, Filing date.
Choose Another Application	Selects a different application, creates a new application,

	creates a template from an existing application (Save As), or deletes an application.
--	---------------------------------------------------------------------------------------

3.1.2 Sequence Functions Data Entry Screens

<i>Data Entry Screen Name</i>	<i>Information Provided</i>
Sequence Data Sequence Editor	Hand key or import sequence data: includes functionality to Add, Modify, or Delete sequences
Organism Information	Source of the nucleic acid or protein.
Feature Information	Areas in the sequence that are biologically significant.
Publication Information	Publications with related sequence information.
Generate Sequence Listing	Create application on disk or paper.
Browse Sequence Listing	Browse, print, or copy to disk the sequence listing.
Help	Accesses on-line help.

3.2 MANEUVERING IN A DATA ENTRY SCREEN

3.2.1 Data Entry Screen Cursor Control Keys

To move to the next field on a data entry screen, use **Enter** or \downarrow (down arrow key). To go to the previous field on a data entry screen, use \uparrow (up arrow key). Pressing **Enter** while in the last field on a data entry screen takes you to the first field. Use \leftarrow and \rightarrow to move to the left or right within a field. Press **Cancel** or **Exit** to exit the data entry screen and return to the *Main Menu/Quick Guide*.

The following table provides a description of the cursor control keys and their corresponding functions.

Go to next field on data entry screen.	Enter or \downarrow
Go to previous field on data entry screen.	\uparrow
Move left within a field.	\leftarrow
Move right within a field.	\rightarrow
Remove the character to the left of the cursor.	Bksp
Remove the character at the cursor.	Del
Save data entry screen and return to <i>Main Menu</i>	Cancel or Exit

3.2.2 Cursor Control Keys for Lists

A data entry window with built-in “down arrows” is called a combination box (or “Combobox”). Comboboxes contain lists. Double clicking on the down-arrow presents the list. Use \downarrow to go to the next item on a list, or \uparrow to go to the previous item on a list. In larger lists (such as the journal names), use the mouse to move within the list. In these large lists, typing a letter scrolls to the first item beginning with that letter. On the Organism Name list, which is a 30,000 record indexed list, scrolling with a mouse is impractical. Therefore, the scroll bar is disabled and you must either page up/down or begin typing the name of the organism to search for it.

Use Enter to select the highlighted item on a list. The program will automatically enter that selection in the appropriate field on the data entry screen. On lists with OK/Cancel options, select Cancel to exit the list without making a choice.

The following table provides a description of the cursor control keys and their corresponding functions.

Go to next item on a list.	-
Go to previous item on a list.	-
Go to next page of a list.	PgDn
Go to previous page of a list.	PgUp
Move cursor to word that begins with a typed character (on longer lists).	A-Z or 0-9
Select the highlighted item on a list and leave the list.	Enter or OK
Leave a list without choosing an item (where applicable).	Exit or Cancel

3.3 PATENTIN V. 2.0 LIMITATIONS

PatentIn v. 2.0 is designed to be as flexible as possible to allow you to enter a wide range of data. However, there are limits on the type and quantity of information PatentIn v. 2.0 can accept.

The limitations existing in PatentIn v. 2.0 are:

- Each sequence is limited to 50,000 residues in the Sequence Editor. If any number of residues exceeds 50,000, whether by hand-keying or importing, the editor will truncate the sequence to the first 50,000 residues.
- There are no programming limits on the number of sequences that you can be added to the sequence list for an application.
- “Publication Title” and “Invention Title” lengths are limited to 200 characters each.
- Feature notes (“Other Information” section) are limited to 200 characters each, with a maximum of one note per feature.

If these limitations restrict you from entering important information for your application, you can use a text editor outside PatentIn v. 2.0 to add information to the Sequence Listing. See **Sections 12.3** and **Section 12.4** for details on editing the Sequence Listing outside PatentIn v. 2.0 and preparing the printed submission.

4. Getting Started - Choose An Application

In Windows 3.1 or 3.11, double click on the icon for PatentIn v. 2.0. If you are using Windows 95 or Windows NT click on the Start button, go to the programs directory, then click on the PatentIn v. 2.0 subdirectory.

*NOTE: This section assumes that the program has been installed according to the instructions found in **Section 2.3** "Installation."*

Select your language preference to be used throughout the program. From the displayed pick list, select from the options: **English, French, German, and Spanish**. The next screen is entitled "**Choose An Application**." It displays a list of applications plus the following functional buttons:

4.1 CREATE

1. If you would like to create a new application (Sequence Listing), click on the **Create** button to create a **New Application**. To open an existing application, skip to **4.2 Open**. When creating a **New Application**, the program prompts you to enter an application name of eight characters or less. Do NOT use any file name extensions! Standard DOS files allow the use of any of the following characters:

A-Z a-z 0-9 \$ % - @ { } ~ ! #

2. After you type a short name, press **Enter**, then enter a "File Reference" name up to 38 characters. If you attempt to exit without entering a "File Reference," the program will ask you to input a "File Reference" in that field. Entries in both fields and a valid file name are required to create a project.
3. When you have finished, click on the Save button. The application will be created. Exit will return you to the **Choose An Application** screen. If you have not saved the application, you will be prompted to do so as you exit.
4. Open the application as directed in 4.2 Open.

4.2 OPEN

If you have already created and saved applications, a list of applications is displayed in alphabetical order. Navigate through the application list by using the arrow keys, the PgUp and PgDn keys, or the mouse to click on the scroll bar. Figure 4-1 shows application list with one application displayed. You can open the desired application by highlighting it and pressing **Enter** or the **Open** button, or by double-clicking on the highlighted file. Proceed to **5.0 Main Menu and Quick Guide** to complete your application.

4.3 SAVE AS

A duplicate of an existing application project can be created as a template by highlighting an application in the list and clicking on **Save As**. Enter a new file name and a new file

reference name in the text boxes, **Save**, then **Exit**. All supporting files and records will be copied from the old application to the newly created application. The old application will be preserved. If you use a standard format when preparing applications, it is recommended that you create a standard application called TEMPLATE with the common fields completed (i.e. Applicant names, title, etc.). This TEMPLATE application can be duplicated with the **Save As** button, and will reduce preparation time.

4.4 DELETE

Delete will remove an application from the list, remove all sequences created, and remove the subdirectory in which they were stored. Deleted applications cannot be restored. If disk storage is limited, it may be necessary to delete old applications.

4.5 EXIT

If no application has been opened, **Exit** closes all files and terminates the program. The system will ask if it is your intention to exit. If not, the exit procedure is aborted, and control returns to the **Choose An Application** screen.

If an application is currently open, **Exit** returns to the Main Menu/Quick Guide, and the current application remains open.

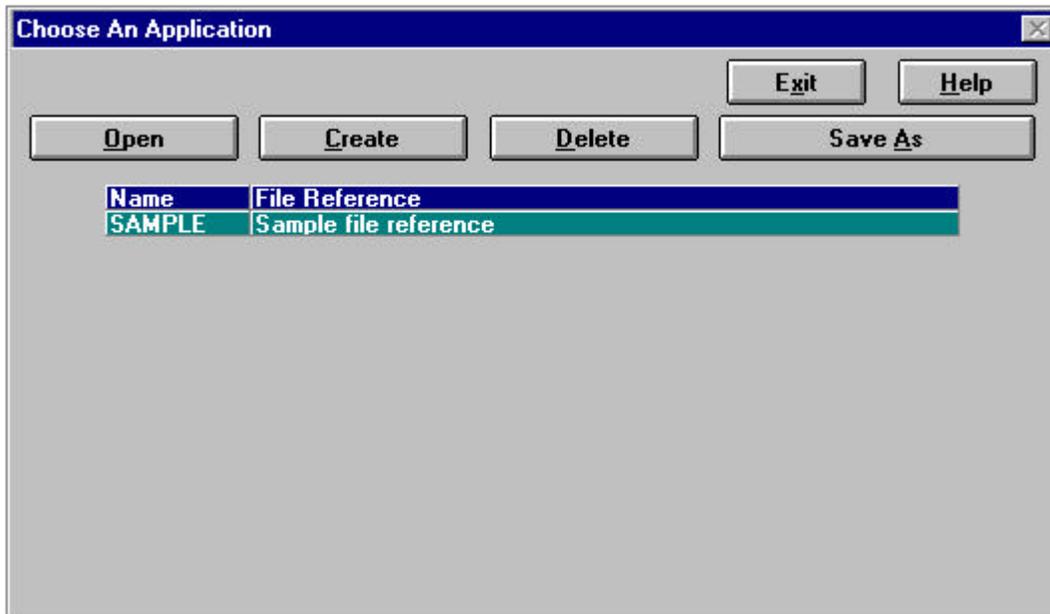


Figure 4-1: Choose An Application

5. The Main Menu and Quick Guide

The *Main Menu/Quick Guide* appears after you create or open an application. (**Figure 5-1** shows the *Main Menu/Quick Guide* screen.) The *Main Menu/Quick Guide* displays the following information on the screen:

- the file name of the current application
- a *Main Menu* bar displaying the key pull-down menus (*Application Data*, *Sequence Functions*, and *Help*), and,
- the *Quick Guide* displays a sequential list of radio buttons that guides the user through the steps in creating a Sequence Listing. (The same steps are contained in the pull-down menu.)

5.1 USING THE MAIN MENU

You select an option by clicking on the menu bar and selecting the function from the pull-down list. When creating a new application, you should use the items on the *Main Menu* in the order they are shown. However, you can complete the data entry screens in any order you prefer.

After you select an option from the *Main Menu*, the program displays the corresponding data entry screen with the blank fields to be completed. Each data entry screen is equipped with field-specific and general help screens that will assist you with the information to enter in the field. Instructions are provided on getting a list of the desired entries. (See **Section 7** “Getting Help” for more information on the help window.)

The Main Menu options are as follows:

5.1.1 Applicant Data

This menu option includes

- Applicant** (See **8.1 Applicant Data Information**)
- Current Application** (See **8.2 Current Application Information**)
- Prior Application** (See **8.3 Prior Application Information**)
- Choose Another Application** (See **4.1 Choose and Application**)

5.1.2 Sequence Functions

This menu option includes

- Sequence Data** (See **9.0 Completing the Sequence Functions**)
- Generate Sequence Listing** (See **12.1 Generating the Sequence Listing**)
- Browse Sequence Listing** (See **12.2 Browse Sequence Listing**)
- Help** – Help screens from the **Sequence Functions** portion of the Help file.

5.1.3 *Exit*

To exit PatentIn v. 2.0, click on the *Exit* at the menu bar, or in the *Main Menu*, click on the **Exit** button. After you confirm that you wish to exit PatentIn Version 2.0, the program returns you to Windows.

5.1.4 *Help*

This help button contains keys to the contents of the entire PatentIn Version 2.0 Help file.

5.2 **THE QUICK GUIDE**

Most users will prefer the *Quick Guide*, a menu displayed in the main window simultaneously with the Main Menu bar. It is organized in functional groups, in the order in which they are normally executed when building an application, and with expanded text. (See **Figure 5-1**.) The *Quick Guide* is more descriptive than the Main Menu. Highlight the button next to the function to be executed, then click the **OK** button to activate the function. The *Quick Guide* buttons operate as follows:

5.2.1 *OK*

The **OK** button activates the menu function for the radio button currently highlighted.

5.2.2 *Cancel*

The **Cancel** button closes the current application and returns to the **Choose An Application** menu. It functions identically to the **Application Data/Choose An Application** selection from the Main Menu bar.

5.2.3 *Exit*

The **Exit** button closes all open files and exits PatentIn Version 2.0. The application asks if this is your intention. If not, the exit routine is canceled. This button functions identically to the **Exit** option on the *Main Menu* bar.

5.2.4 *Help*

This is the main **Help** function. It displays the Contents of the PatentIn Version 2.0 Help file. It calls the same Help screen as does the **Help** option on the *Main Menu* bar.

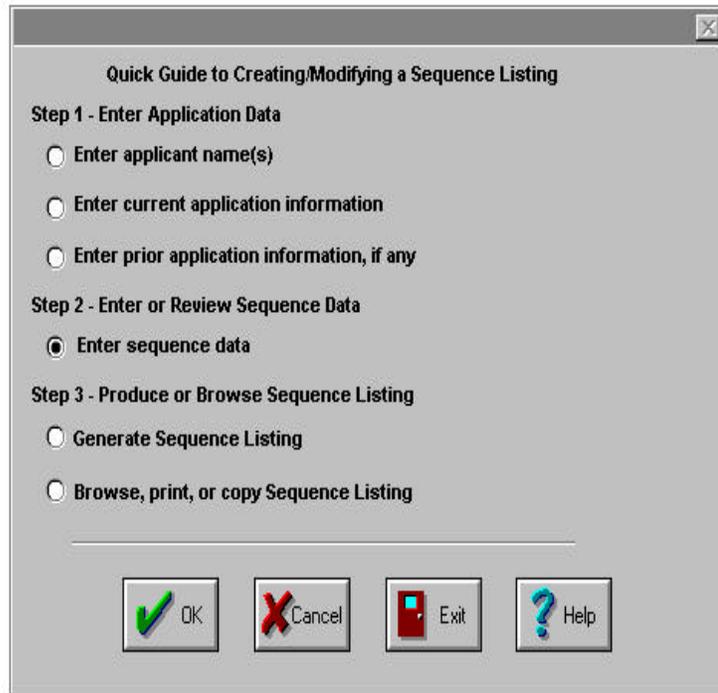


Figure 5-1: PatentIn v. 2.0 Main Menu/Quick Guide

6. Tips for Completing Data Entry Screens

6.1 COMPLETING FIELDS IN DATA ENTRY SCREENS

You can complete most fields by either typing information directly into the field or selecting from a list of choices.

In most cases, a field that is filled from a list is a Combobox (See “Combination box” in **2.2.1 Definitions**). Click on the down-arrow to the right of the combobox to open the list. Click on the selection to fill the box and hide the list.

Some lists that are very large (e.g. the Organism Name list), or that are associated with multi-line fields (e.g. the Additional Information lists) are displayed in a listbox with control buttons and a specialized search feature.

Data for most functions is collected in one screen. The exception is the **Publications** data entry function (under the **Sequence Functions** option of the *Main Menu* or **Enter Sequence Data** option of the *Quick Guide*.) There are three Publications related data entry screens, in addition to associated lists. In these cases, the buttons **Next** and **Previous** transfer you to separate sub-screens where you type the information or make selections from lists.

6.2 MANDATORY FIELDS IN DATA ENTRY SCREENS

PatentIn v. 2.0 verifies that the following required fields are completed before compiling the Sequence Listing:

Applicant Information Data Entry Screen

<110> • **Applicant Last Name, First Name, or Organization Name**

Current Application Information Data Entry Screen

<120> • **Invention Title**

<130> • **File Reference**

<140> • **Application Number (if known)**

<141> • **Filing Date (if known)**

Sequence Editor

<213> • **Organism**

<400> • **Sequence**

Mandatory Information Automatically Supplied by PatentIn for Windows

<160> • **Number of sequences**

<170> • **Software**

<210> • **Information for SEQ ID NO**

<211> • **Length (of sequences)**

<212> • **Type (of sequence: DNA, RNA, or PRT)**

6.3 LISTS WITH RESTRICTED VOCABULARY

PatentIn v. 2.0 provides lists from which you can select values and options to enter into fields. Some lists are editable, while others are restricted.

In the Publication Name list in the Bibliographic Information screen, if the value you want is not on the list, you may type **Other** in the field. This calls a data entry screen where a journal name can be entered, which can be added permanently to the journal list for future reference. While the Organism Name list cannot be edited directly, the Organism menu has an “**Other**” option, which will permanently add your new organism name to the list.

The list of feature (Feature Name/Key) is restricted to the values from tables 5 and 6 in WIPO ST. 25. (See **Chapter 13**.)

6.4 SAVE AND EXIT BUTTON BEHAVIOR THROUGHOUT PATENTIN VERSION 2.0

Save and **Exit** button behavior throughout PatentIn is conventional except for the *Library of Previously Defined Publications* in **Section 9.1.6**. Any selection made within the *Library of Previously Defined Publications* will be automatically saved when the **Exit** button is clicked.

7. Getting Help

7.1 LEVELS OF HELP AVAILABLE

There are three levels of help available in PatentIn v. 2.0:

- General help
- Field specific help for particular fields within a screen
- Data Entry Screen specific help

7.1.1 General Help

General help is obtained by clicking the *Help* option from the *Main Menu*, from the menu bar, or the Help button on the Quick Guide. The file displays information about the use of PatentIn v. 2.0 and about patent-related topics.

Click on the green hypertext to systematically review all help topics.

7.1.2 Field Specific Help

Located at the bottom of the screen is the status bar, which provides context specific instructions based on the placement of the cursor in any field. To find out what information a field requires, simply move the mouse indicator over that particular field and the instructions will appear in the status bar.

7.1.3 Data Entry Screen Help

Every data entry screen is also programmed with a Help button. Click on this button to access context specific help.

8. Completing the Application Data

This section describes the functions for completing the Application Data.

8.1 APPLICANT INFORMATION

8.1.1 Applicant Information Data Entry Screen

Selecting “Enter Applicant Name(s)” from the *Quick Guide* initially presents the Applicant Information data entry screen if there are no applicants. When applicants are present, the “Applicant List” browse screen is presented, containing a list of all applicants. From this screen you have the option to Modify, Add, or Delete an applicant record.

The Applicant data screen that follows is used to enter the names of the applicants, as well as optional correspondence information for the application. In the Applicant Information data entry screen, you must first indicate whether the applicant is an Individual or Organization. If you select Individual, which is the default option, the Applicant Information data entry screen then allows you to enter the name(s) (Last Name, Suffix, First Name, and Middle Initial) of the person(s) applying for the patent. Choosing Organization allows you to enter the Organization Name.

“First Name” and “Last Name” for Individual, and “Organization Name” for Organization are mandatory entries. Correspondingly, the text for these fields is shown in the color red.

You may also enter optional information in the remaining fields (Address, City, State, ZIP, Phone Number, Fax Number and E-mail). These fields are for your internal user records and reference.

When you have finished entering the applicant information, click either the Save or Exit buttons. If you click on the Save button, the information entered is saved to the current application. If you click on the **Exit** button without saving, the program will prompt you to save the entered information. If you elect not to save your data and there are no applicants, you will be returned to the *Main Menu/Quick Guide*. If there are any applicants, you are returned to the *Applicant List* browse screen. Exiting the *Applicant List* returns you to the *Main Menu/Quick Guide*. **Figure 8-1** shows an example of the "Applicant Data" data entry screen.

Figure 8-1: Applicant Information Data Entry Screen

8.1.2 Applicant Information Browse Screen

After saving the information in the Applicant Data Entry Screen, you automatically access the Applicant Information browse screen. **Figure 8-2** shows an example of the *Applicant List* browse screen. This browse screen is a list of all applicants entered for the current application. While in this browse screen, your options are:

- **Modify** - You may edit the information on a particular applicant by highlighting the desired applicant and clicking on the **Modify** button.
- **Add** – You may add more applicants to the application by clicking the **Add** button and entering the appropriate information.
- **Delete** – You may delete an applicant by highlighting the particular applicant and clicking the **Delete** button.
- **Exit** – After entering all applicants, press the **Exit** button from the browse screen to return to the *Main Menu/Quick Guide*.

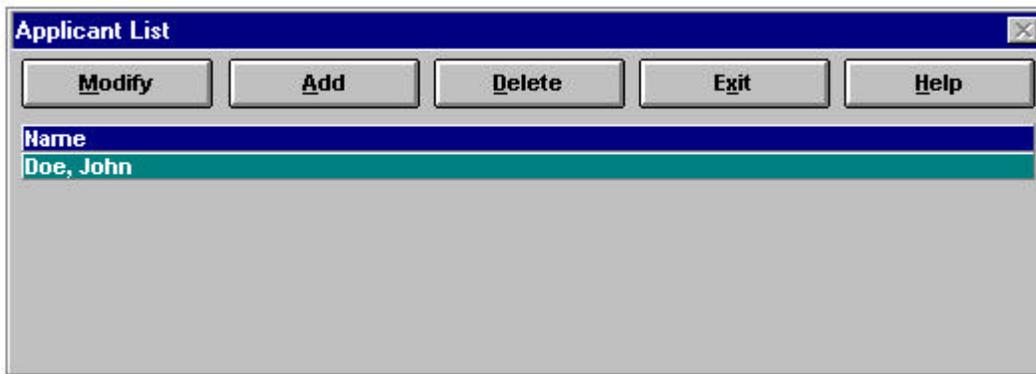


Figure 8-2: Applicant Information Browse Screen

8.2 CURRENT APPLICATION INFORMATION

8.2.1 Current Application Data Entry Screen

Selecting “Enter Current Application Information” from the *Quick Guide* presents the Current Application data entry screen. This screen collects information about the application that is currently being filed. The fields in this screen are: **Title**, **Application Number**, **Filing Date**, and **File Reference**. These fields collect required information about the current application. Figure 8-3 shows an example of the Current Application Data Entry Screen.

*NOTE: If you submit the Sequence Listing after filing the original application, such as an amendment, you must provide the **Application Number** and **Filing Date**.*

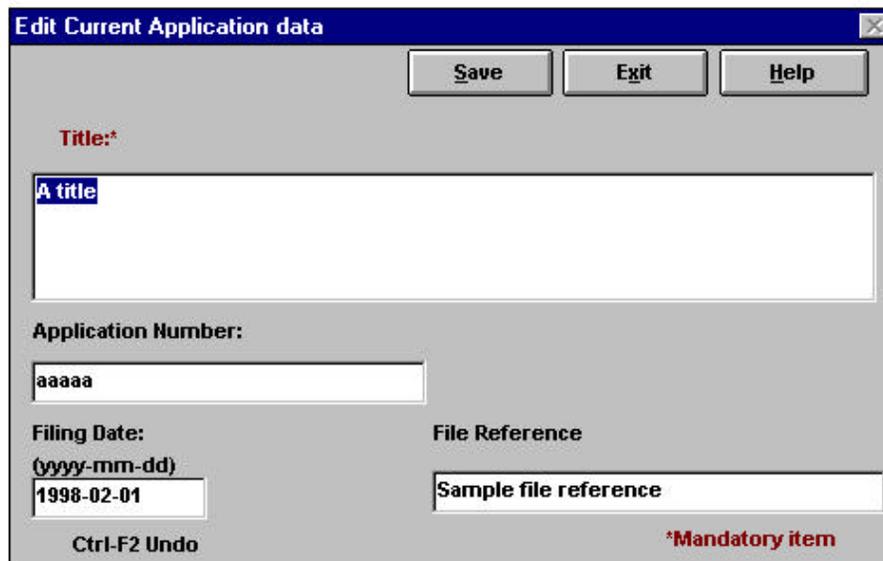


Figure 8-3: Current Application Data Entry Screen

8.2.1.1 Title

Enter a title for the application in this field. This is a mandatory field.

A 200-character field is displayed for entering, reviewing, and editing the title. To leave the title field, press the **Tab** key, and the cursor will move to the next field.

8.2.1.2 Application Number

If known, enter the application number in the format that is required for the receiving office you are submitting your sequence listing to. If you are unsure of the required format for this field, refer to WIPO ST.10/C for the proper format to use. PatentIn will *not* check the Application number field for correct format.

NOTE: *The Application Number field is not mandatory, but if completed, you must complete both the Application Number and Filing Date fields.*

For PCT applications, type the complete PCT application number including the letters PCT, a slash, a two-letter code indicating the Receiving Office, a two-digit indication of the year, a slash, and a five-digit number in the Application Number field; for example, PCT/US88/99999.

8.2.1.3 Filing Date

Enter the original filing date of the application in this field, if known. Enter the date in the format **YYYY-MM-DD**, for example, 1997-12-31.

NOTE: *The Application Number field is not mandatory, but if completed, you must complete both the **Application Number** and **Filing Date** fields.*

8.2.1.4 File Reference

This is a description of the application used as a file reference. It may be a maximum of 38 characters and spaces. The data that was entered in the “File Reference” field when you created your project will be shown here. The file reference name can be edited at any time by using this screen.

When you have finished entering the application information, click either the Save or Exit buttons. If you click on the Save button, the information entered is saved to the current application. If you click on the **Exit** button, the program will prompt you to save the entered information, then returns you to the *Main Menu/Quick Guide*.

8.3 PRIOR APPLICATION INFORMATION

8.3.1 Prior Application Information Data Entry Screen

Selecting “Enter Prior Application Information” from the *Main Menu* for the first time, displays the Prior Application Information data entry screen. This screen collects information about any previous domestic, foreign priority, or international patent applications pertaining to the sequence currently being patented.

If any records are present, the first screen will be a list of these records. From this list you may Modify, Add, or Delete a Prior Application record.

The fields to be completed in this screen are the Application Number and Filing Date.

Figure 8-4 shows an example of a Prior Application Information Data Entry Screen.

NOTE: The Prior Application data entry screen is not mandatory, but if filled out, both the Filing Date and Application Number fields must be completed.

Figure 8-4: Prior Application Data Entry Screen

8.3.1.1 Application Number

If known, enter the application number in the format that is required by the receiving office to which you are submitting your sequence listing. If you are unsure of the required format for this field, refer to WIPO ST.10/C for the proper format to use. PatentIn will not check the Application number field for correct format.

NOTE: The Application Number field is not mandatory, but if completed, you must complete both the Application Number and Filing Date fields.

If you are entering a PCT application, type the complete PCT application number, including the letters **PCT**, a slash (/), a two-letter code indicating the Receiving Office, a two-digit indication of the year, a slash, and a five-digit number in the **Application Number** field; for example, **PCT/US88/99999**.

8.3.1.2 Filing Date

Enter the date on which the prior application was filed. Enter the Filing Date in the format YYYY-MM-DD, for example, 1997-12-31.

NOTE: The Application Number field is not mandatory, but if completed, you must complete both the Application Number and Filing Date fields.

When you have finished entering the applicant information, click either the Save or Exit buttons. If you click on the Save button, the information entered is saved to the current application. If you click on the **Exit** button, the program will prompt you to save the entered information, then opens the Prior Application browse screen if there are any Prior Application records. Otherwise, PatentIn returns to the *Quick Guide*.

8.3.2 Prior Application Information Browse Screen

After saving the information entered in the Prior Application data entry screen, you are returned to the Prior Application Information browse screen which displays a list of all prior applications relevant for the current application. **Figure 8-5** shows an example of the Prior Application Information Browse Screen. While in this browse screen, you have the following options:

- **Modify** – You may edit the information on a particular prior application by highlighting the desired applicant and clicking on the **Modify** button.
- **Add** – You may add prior applications to the current application by clicking the **Add** button and entering the appropriate information.
- **Delete** – You may delete a prior application by highlighting the particular applicant and clicking the **Delete** button.
- **Exit** – After entering all of the applicants, press the **Exit** button from the browse screen to return to the *Main Menu/Quick Guide*.

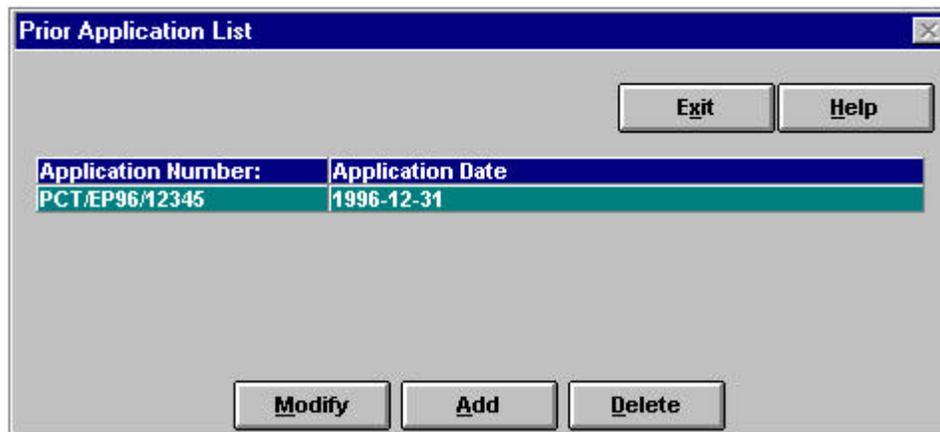


Figure 8-5: Prior Application Information Browse Screen

9. Completing the Sequence Functions

The *Sequence Functions* menu options enable you to enter data relevant to your application. Selecting “Enter Sequence Data” from the Quick Guide accesses the Sequence Data browse screen. From this screen, entitled “Sequence Functions for Current Highlighted Sequence,” you may access the supporting functions for these sequences.

These supporting functions include buttons to open the other data entry screens for Organism, Feature, and Publications. They also include buttons to **Modify** a sequence on the list, **Add** a new sequence in the Sequence Editor, **Import** a sequence to the list from another directory, **Delete** a sequence from the list, or **Restore** a previously deleted sequence to the list.

9.1 SEQUENCE DATA BROWSE SCREEN

Select “Enter Sequence Data” from the *Quick Guide* to access the Sequence Data browse screen. **Figure 9-1** shows an example of the Sequence Data Browse Screen. The Sequence Data browse screen lists the file name, type, and length of sequence, current numerical order of sequences, and whether an Organism – a mandatory field – has been assigned to the sequence.

The order of sequences may be changed at any time by “drop and drag.” (For more information, see **Section 10.4** “Reordering Sequences.”) The order of sequences is renumbered automatically, and therefore always appears in ascending order. The order number may not be consecutive if there are deleted sequences, since deleted sequences retain their order number.

The order numbers correspond to the order in which the sequences are evaluated when compiling the sequence listing. However, this may not necessarily be the sequence number itself, since sequence numbers are assigned at the time the sequence is compiled based on visible sequences only. Furthermore, CDS features insert additional sequences with their own sequence numbers. (See **Section 11.9** “Annotating a Protein Coding Region” for additional information.)

This screen also provides important links to other screens, such as **Organism Information, Feature Information, Publications Selections Browse Screen, Import, Restore**, and the **Sequence Editor**. These screens, with the exception of the Sequence Editor and Feature Information, are discussed in the following sections.

NOTE: For details on the Sequence Editor, refer to **Section 10** “Using the Sequence Editor.” For details on the Feature Information, refer to **Section 11** “Completing the Feature Information.”

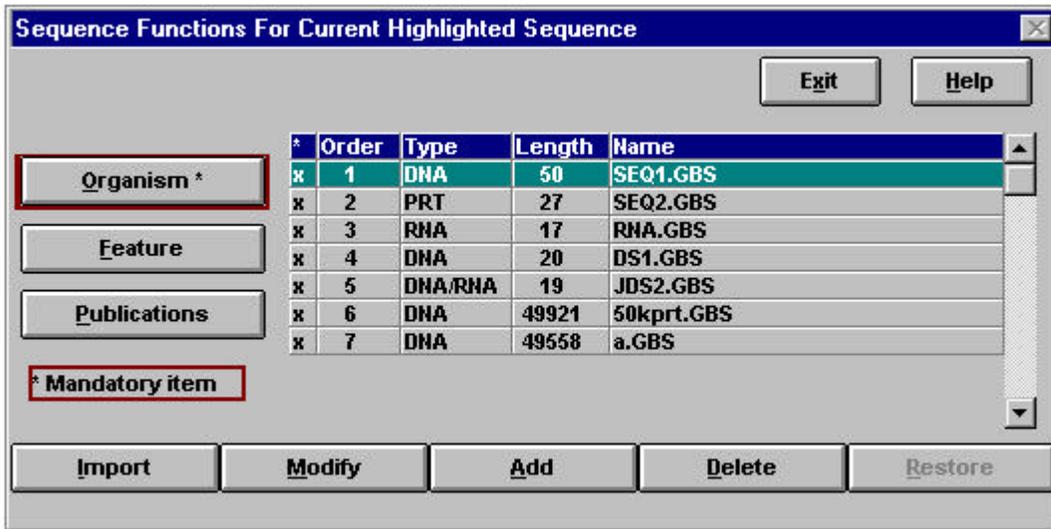


Figure 9-1: Sequence Data Browse Screen

9.1.1 Organism

To assign an organism to a sequence, first highlight the desired sequence, then click on the **Organism** button from the Sequence Data browse screen to access an intermediate screen. In accordance with WIPO ST.25, each sequence must specify the original source organism. This screen displays a list box enabling you to select from the radio-button options of **Organism**, **Unknown**, **Other**, and **Artificial Sequence**. **Figure 9-2** shows an example of the Organism Button selections.

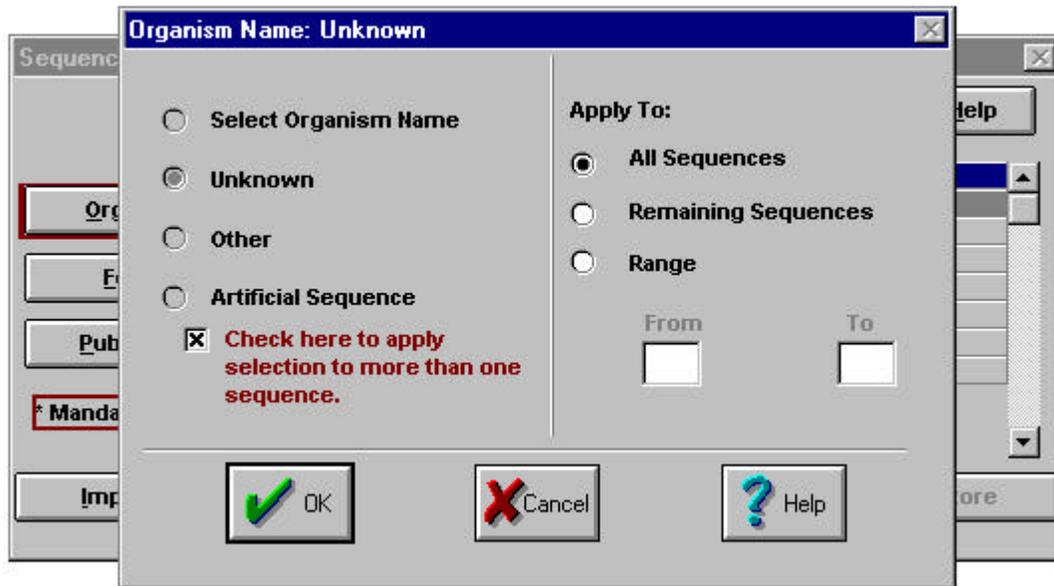


Figure 9-2: Organism Button Options

Select Organism Name opens a list of selected known Organism names. You locate your selections by typing the organism name in the text box below the list or by using the cursor control keys: page down/up or arrow down/up. Do not mouse-click in the text box first. Just immediately start typing the first characters of the organism name. If your desired organism name is not in the list, click on cancel and select “Other” within the main organism menu.

If you select **Unknown**, you will also be presented with a data entry box to describe the Unknown organism further. This entry is mandatory as per WIPO ST. 25. It appears as a Feature note associated with this sequence, and can be further modified in the Feature section.

Select the **Other** radio button if the organism you require is not indicated on the “Select Organism Name list.” A dialog box will open, allowing you to enter the appropriate scientific name of the organism. WIPO ST.25 dictates that the response for the organism field be the Latin species name. The user will be asked to type in the name twice if the new name will be permanently added to the “Organism” list.

If you select **Artificial Sequence**, you will also be presented a data entry box to describe the Artificial Sequence further. This entry is mandatory. It appears as a Feature note associated with this sequence, and can be further modified in the Feature section.

9.1.2 Feature

Click on the **Feature** button to access the Feature Information data entry screen. (For detailed information about this screen, refer to **Section 11** “Completing the Feature Information.”)

9.1.3 Publications

PatentIn v. 2.0 allows you to enter as many references relating to your sequence as needed. These publications are entered on the Bibliographic Information data entry screens and author screens, which are entered through the *Library of Previously Defined Publications* Bibliographic Information Data Entry Screen (See **Figure 9-3**).

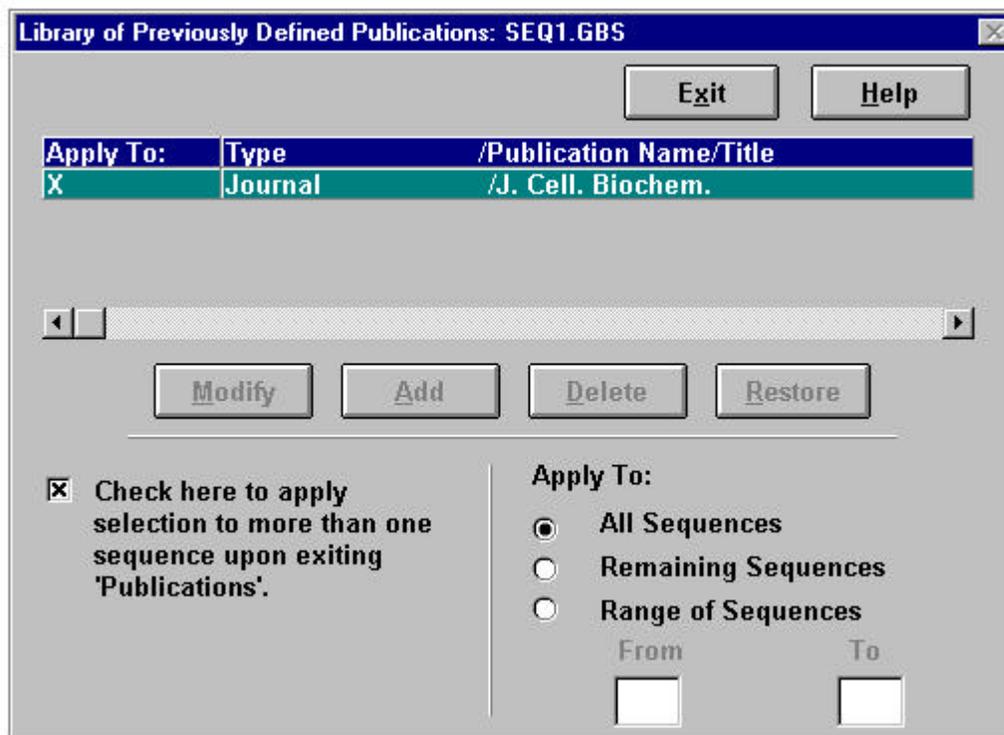


Figure 9-3: Library of Previously Defined Publication

Activate **Add** to open the first Bibliographic Information data entry screen. Once you have completed the screen's fields, click on the **Next** button to access additional fields and screens. The third screen is a list of **Authors**, from which individual author records may be selected to **Modify**, or a blank screen opened to **Add** authors.

Each field on the Bibliographic Information data entry screen is described in the following sections. **Figure 9-4** shows an example of the Bibliographic Information Data Entry Screen.

Figure 9-4: Bibliographic Information Data Entry Screen

9.1.3.1 Publication Type

For this field, you must select one of the following publication types from a list: **Journal**, **Patent**, or **Thesis**. Clicking on the field enables you to select your entry choice.

<i>Publication Type</i>	<i>Description</i>
Journal	Journal publication. The Journal category includes journals (such as <i>Cell</i> or <i>Gene</i>) as well as serial publications (such as <i>Cold Spring Harbor Symposia</i> volumes and <i>Methods in Enzymology</i>). A list of approximately 300 journals is available under the “Journal Name” option, and the “Other” selection permits adding another. This field also permits entering a database accession number and database entry date. <i>NOTE: If the sequence will eventually appear in a journal yet to be published, choose Journal. You can then indicate the sequence's pre-publication state in the next field.</i>
Patent/Published Application	Patent/Application publication. This selection changes the fields in the first Bibliographic Information screen to the Publication Database Entries screen. Journal information is removed and a pop-up data entry box is presented to enter a published application's number, filing date, and publication date.
Thesis	Thesis publication. This selection presents the same data entry fields as Journal, except the database accession number, database entry date, and page number references are omitted.

9.1.3.2 Publication Name

If the specified sequence appears in a journal, click in the Publication Name field and select the appropriate journal name from the pick list of approximately 300 journals. This list is alphabetically indexed. To find the associated journal, click in the field to display a scrolling list of the journals from which to make your selection. A journal can also be located by first-letter keystroke (i.e., pressing the letter “U” calls the first journal beginning with “U.” Pressing “U” again skips to the next item beginning with “U.” etc.).

After you make a selection, the journal name appears in the field. If the journal is not on the list, select “Other” from the list in order to access a data entry dialog box. In this box, type the name or official abbreviation of the journal that will appear in the Publication Name field. The name needs to be typed in twice to ensure spelling if it will be permanently added to the Journal list.

9.1.3.3 Publication Date

Enter the year or the season in which the article, journal, serial, or thesis was published in the **Publication Date** field. If this information is unknown, skip this field. This is NOT a standard date field and does not need to be completed in the conventional date format.

9.1.3.4 Volume and Issue

Enter the volume of the journal or serial in the **Volume** field. If the Publication Information includes an issue number, enter it in the **Issue** field.

9.1.3.5 Start Page and End Page

In the appropriate field, enter the page on which the Publication Information starts and ends. In the **End Page** field, include the pages containing references. You do not need to enter page numbers for a thesis, and therefore these fields are not visible.

***NOTE:** At this point in the screen, you will need to click on the **Next** button to access the remaining fields. See **Figure 9-5** for an example of the Bibliographic Information Data Entry Screen, continued.*

9.1.3.6 Title

This field should contain one of the following: Journal Article Title, Serial Article Title, or Thesis Title.

Because the Publication Title is a lengthy field, PatentIn v. 2.0 provides a simple mini-editor for entering, reviewing, and editing. The entered text is considered a single “line,” and the editor automatically wraps and displays it, up to 4 lines at a time, on the screen. The **Title** field is limited to 200 characters of text. When you are finished typing, press either the **Tab** or **Enter** key to move to the next field.

9.1.3.7 Relevant Residues (From/To Fields)

These fields indicate the part of the sequence, or “residues,” reported in a particular publication. In many cases, the entire sequence is included in a publication; however, in other cases, only a portion of the sequence is included in a publication. In the **From** field,

enter the first residue of your application sequence reported in the publication you are entering. In the To field, enter the last residue of your application sequence reported in this publication.

If the numbering scheme used for the residues in the publication differs from that used by PatentIn v. 2.0, the values that you entered in the **From** and **To** fields should reflect the numbering as it appears in the PatentIn v. 2.0 Sequence Editor. The PatentIn v. 2.0 Sequence Editor always begins numbering at "1" and rejects negative numbers.

Figure 9-5: Bibliographic Information, continued Data Entry Screen

9.1.4 Author Information Browse Screen

Clicking on the Next button accesses the Author Information browse screen. Figure 9-6 shows an example of the Author Information Browse Screen. The authors of the publication are displayed in a list. Unlike comparable lists in PatentIn, this list appears whether or not authors are present. The options in this screen are:

- **Modify** - You may edit the information on a particular author by highlighting the desired author name and clicking on the **Modify** button.
- **Add** – You may add more authors to the application by clicking the **Add** button and entering the appropriate information.
- **Delete** – You may delete an author by highlighting the particular author name and clicking the **Delete** button.
- **Exit** – After entering all authors, press the **Exit** button from the browse screen to return the “Library of Previously Defined publications.”

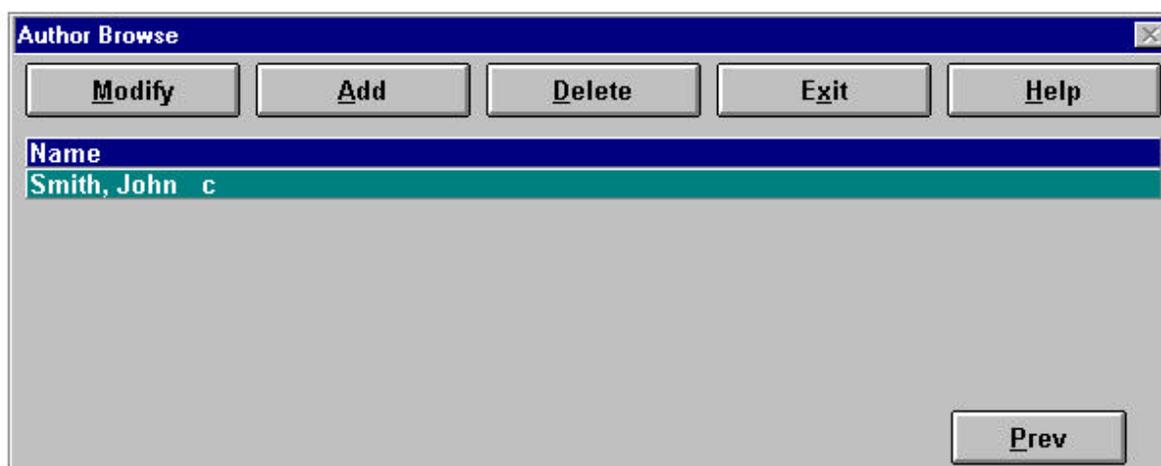


Figure 9-6: Author Information Browse Screen

9.1.5 Author Information Data Entry Screen

If you select **Modify** or **Add** from the Author Information browse screen, you access the Author Information data entry screen. In this screen, enter the names of the authors of the Bibliographic Information in the order in which they appeared (or will appear) in the publication.

When you have finished entering the author information, click either the Save or Exit buttons. If you click on the **Save** button, the information entered is saved to the current application. If you click on the **Exit** button, the program will prompt you to save the entered information, then returns you to the Bibliographic Information browse screen.

9.1.6 Library of Previously Defined Publications Browse Screen

Click on the **Publications** button to access the *Library of Previously Defined Publications* browse screen, if any are present. This screen lists all publications with information relating to the current application that have been created, and permits selected publications to be applied to the current sequence (identified in the title bar). This is done by double-clicking on the record, which toggles the left column in the record with an “X” or a blank to mark and unmark the record. A check box and radio menu options are presented below the function buttons to apply these selections to other sequences as well.

This screen serves as a master list of all references for the entire sequence listing and the individual selections for the current sequence under consideration, and provides tools to expedite the process of selecting publications. Figure 9-3 shows an example of a *Library of Previously Defined Publications* browse screen.

After saving all the information entered in the Author Information data entry screen and browse screen, you are returned to the initial library browse screen that lists all

publications and publication types entered. **Figure 9-3** shows an example of the Library of Previously Defined Publications. While in this browse screen your options are:

- **Modify** – You may edit the information on a particular publication by highlighting the desired applicant and clicking on the Modify button.
- **Add** – You may add more publications to the application by clicking the Add button and entering the appropriate information.
- **Delete** – You may delete a publication by highlighting the particular applicant and clicking the Delete button.
- **Restore** – You may undo all previous deletions in this screen.
- **Exit** – After entering all publications, press the Exit button from the browse screen to save data and return to the Main Menu/Quick Guide.

To apply the selections to other sequences mark the check-box that states, “Check here to apply selection to more than one sequence upon exiting.” That enables the radio menu of the following choices:

- **All Sequences** – All sequences current in the list will reflect the same publications selections. However, sequences that are subsequently added will not be effected.
- **Remaining Sequences** – All sequences after the current sequence in the sequence list will reflect the same publications information. Each sequence will retain its publications selection even if the order is subsequently changed.
- **Range of Sequences** – A specific range of sequences can be specified, based on the current order. There are positions to enter the starting (**From**) and ending (**To**) sequence order numbers. Each sequence will retain its publications selection even if the order is subsequently changed.

The selections are applied when the screen is exited. (See **Figure 9-3**.)

9.1.7 Import

Click on the **Import** button to import ASCII sequence files located on a diskette, hard drive, or file server. Each imported sequence must be within a separate file. However, PatentIn can “bulk” import if the user selects multiple files. The file size is displayed with the file name in the Files window to facilitate identifying files to import. Figure 9.7 shows an example of the Import File Manager. PatentIn can import any text file. Word processing and database files cannot be read unless first stored as text files.

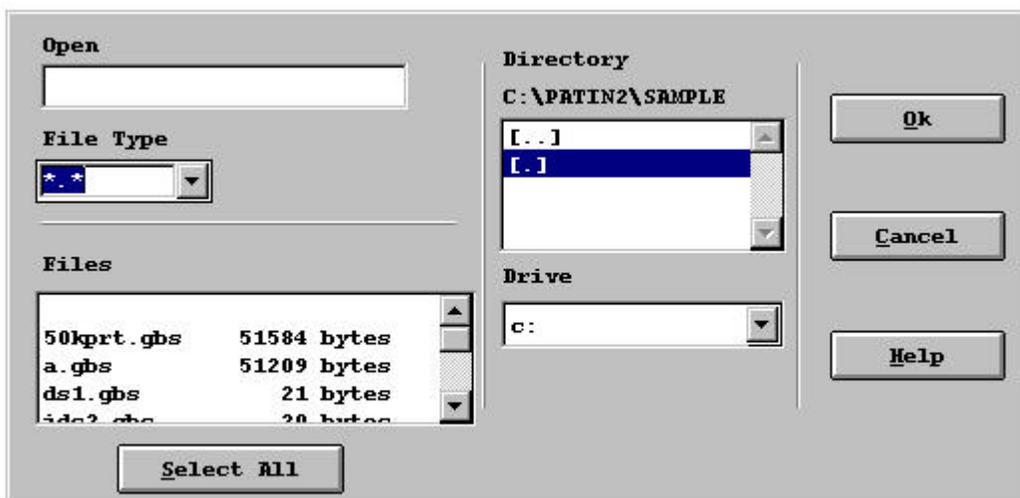


Figure 9-7: Import File Manager

Import has five filters to select from: **DNA Filter**, **RNA Filter**, **Combined DNA/RNA**, **PRT One Letter Filter**, and **PRT Three Letter Filter**. When you open a non-PatentIn v. 2.0 file, the program removes all characters except valid sequence characters. PatentIn v. 2.0 also removes all invalid characters such as line breaks, spaces, and numbers, etc. when opening the sequence, and saves the sequence in PatentIn v. 2.0 format with a .GBS extension. The following characters remain in the sequence after filtering (characters may be in either upper or lowercase):

DNA Filter	a, g, c, t, y, r, m, k, s, w, h, b, v, d, n
RNA Filter	a, g, c, u, y, r, m, k, s, w, h, b, v, d, n
Combined DNA/RNA Filter	a, g, c, u, r, y, m, k, s, w, b, d, h, v, n, t
PRT One Letter Protein Filter	A, B, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, X, Y, Z
PRT Three Letter Protein Filter	Refer to Section 10.3.3.5 “Amino Acid Codes Table”

NOTE: The **RNA and DNA Filters** do not perform any conversions when you import the sequence. For example, if you select the **RNA Filter** and there are **T's** in the sequence, they will be filtered out, not converted to **U's**. The **PRT One-Letter Filter** will also remove all characters that are not standard one-letter amino acid characters (see

*Sections 9.3.3.4 and 9.3.3.5). **PRT Three-Letter Filter** will read valid three-letter amino acid groups and convert them to **PRT One Letter format**. Three-letter groups can be in upper or lower case, and may be separated by a space character. When the Sequence Listing is printed, amino acid residues will be shown in the required three-letter format.*

You can import sequence files that have been created with other editors, word processing programs or data analysis programs provided they have been stored as text files. For sequences that will be imported into PatentIn v. 2.0, be sure to specify a text mode when you save a sequence file from a word processor.

9.1.8 Modify

The **Modify** button opens a previously entered sequence into the Sequence Editor. Highlight the desired sequence and activate the button, double click on the sequence record, or press **Enter**. See **Chapter 10** for detailed instruction on using the Sequence Editor.

9.1.9 Add

The **Add** button opens the Sequence Editor to create a new sequence (by hand-keying). The new sequence is not saved to the project file until saved within the Sequence Editor. See **Chapter 10** for detailed instruction on using the Sequence Editor.

9.1.10 Delete

A sequence record may be deleted by highlighting the target file and activating **Delete**. This hides the sequence and the associated **Organism, Feature, and Publications** selections. These files remain resident within PatentIn v. 2.0 and may be recalled. See **Section 9.1.11 "Restore"** on procedures to recall a deleted sequence. The sequence files are destroyed only when an Application Project is deleted from the *Choose An Application* screen.

9.1.11 Restore

A deleted sequence may be restored using the **Restore** button. When the button is activated, it presents a list of all sequence files that have been deleted for this Application Project. Highlight the sequence to restore and select *OK*. Only one sequence can be restored at a time. The new sequence is restored in the same position that it was in before being deleted, along with all of its associated **Organism, Feature, and Publications** selections.

10. Using the Sequence Editor

10.1 SEQUENCE EDITOR SCREEN

The Sequence Entry Window displays bases of nucleic acid or residues of protein sequence entries. The design of the Sequence Editor is different from the Application data entry screens. The Sequence Editor screen is divided into two parts: a **Menu Bar** including a *Help* option at the top, and a **Sequence Entry Window** encompassing most of the screen. The menu bar provides access to three menus: **File, Attributes, and Search**, which are described in **Section 10.3** “Commands Available from the Menu Bar.”

There are two ways to enter a sequence: type the sequence directly into the Sequence Entry Window, or open a sequence from another file into the editor. Opening files into the editor is described in the File Menu section below under the **Open** option (see **Section 10.3.1.2** “Open”).

Additionally, the Sequence Entry Window displays residues in groups of ten, with sixty residues on each line. The rightmost column of the window displays the number of the last residue on that line. The sequence length and cursor positions are displayed at the bottom of the Sequence Editor screen. **Figure 10-1** shows an example of the Sequence Editor screen.

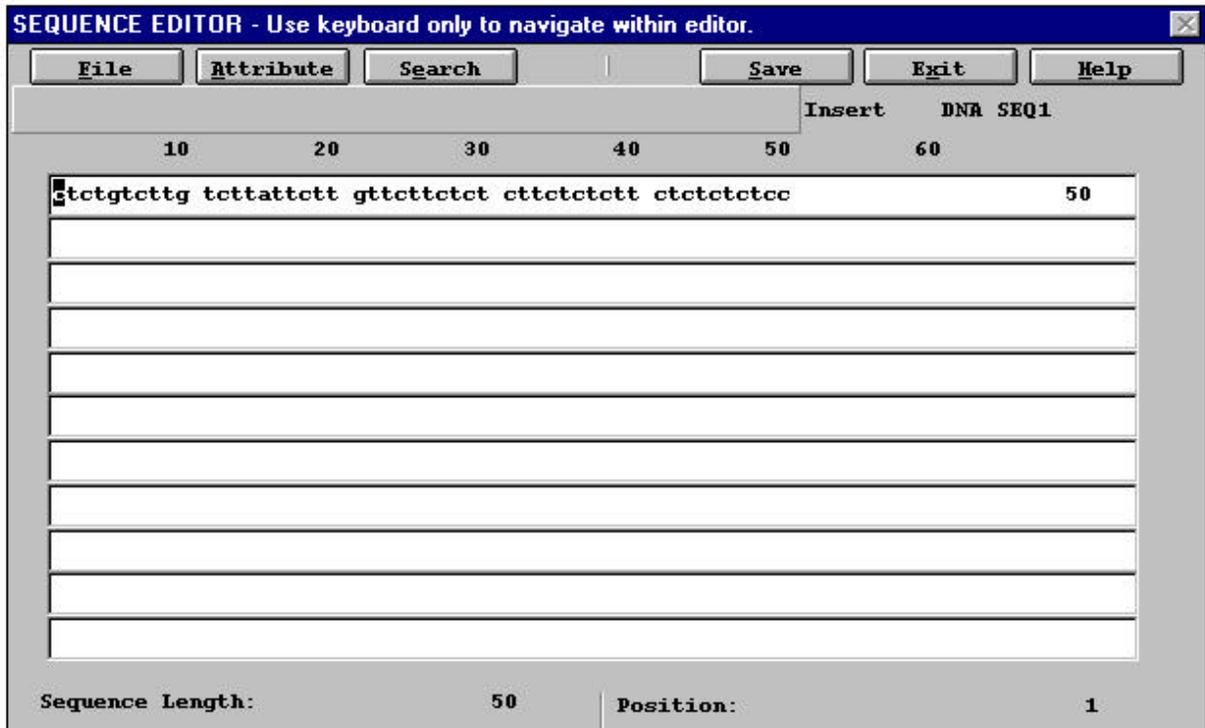


Figure 10-1: Sequence Editor

10.2 ACCESSING THE SEQUENCE EDITOR

To edit a sequence in the Sequence List in the Sequence Data browse screen, highlight the sequence and click on the **Modify** button, or double click on the desired sequence. This will open that sequence within the Sequence Editor.

To create a new sequence, activate the **Add** button. A radio menu will ask for the Sequence Type (the default is **DNA**). After designating the sequence type, the *Sequence Editor* will open with the appropriate filter key set for the operator to type in a new sequence.

At the top of the Sequence Editor is the menu bar with the options **File**, **Attributes**, and **Search**. Click on any of the options on the Menu Bar to activate that particular option. Press **Cancel** to leave the option that you just chose from the Menu Bar.

10.3 COMMANDS AVAILABLE FROM THE MENU BAR

The following sections explain the functions of the Sequence Editor menu bar and the different options available.

10.3.1 FILE Menu

The commands in this menu perform various file-related functions, or allow you to exit the Sequence Editor. The commands in this menu are **New**, **Open**, **Save**, **Save As**, and **Cancel**. Each command is discussed below.

10.3.1.1 New

New allows you to create a new sequence.

10.3.1.2 Open

Use the **Open** command to open existing sequence files into the editor to build a new sequence file. The sequences must be in text only (or ASCII) format. In addition, each sequence file should contain only one sequence; the **Open** command cannot separate multiple sequences contained in one file. However, PatentIn v. 2.0 can import both PatentIn and non-PatentIn formatted text files. Any invalid characters in the imported text file will be filtered out.

Open has five filters to select from: **DNA Filter**, **RNA Filter**, **Combined DNA/RNA**, **PRT One-Letter Filter**, and **PRT Three-Letter Filter**. When you open a non-PatentIn v. 2.0 file, the program removes all characters except valid sequence characters. PatentIn v. 2.0 also removes all invalid characters such as line breaks, spaces, and numbers, etc. when opening the sequence. The sequence becomes part of whatever sequence is in the editor. The following characters remain in the sequence after filtering (characters may be in either upper or lowercase, depending on sequence type):

DNA Filter	a, g, c, t, y, r, m, k, s, w, h, b, v, d, n
RNA Filter	a, g, c, u, y, r, m, k, s, w, h, b, v, d, n
Combined DNA/RNA Filter	a, g, c, u, r, y, m, k, s, w, b, d, h, v, n, t
PRT Filter	A, B, C, D, E, F, G, H, I, K, L, M, N, P, Q, R, S, T, V, W, X, Y, Z

NOTE: The **RNA** and **DNA Filters** do not perform any conversions when you import the sequence. For example, if you select the RNA Filter and there are T's in the sequence, they will be filtered out, not converted to U's. The **PRT One-Letter Filter** will also remove all characters that are not standard one-letter amino acid characters (see **Sections 10.3.3. and 10.3.3.5**). The **PRT Three-Letter Filter** will convert any valid three letter grouping into its one letter counterpart as it is brought into the editor.

WARNING: If the sequence characters shown in the previous table appear anywhere in the file except the sequence, you must use a text editor outside PatentIn v. 2.0 to delete them before importing the sequence. Otherwise, PatentIn will import the character as a codon. Be sure to check that the length of the imported sequence is as expected. The maximum length of any sequence file can be only 50,000 characters. The source file may be any length so long as the number of valid codons does not exceed 50,000.

To open a sequence into the Sequence Editor, click on Open. You will access the import file manager, in which you indicate the drive, directory, and file name of the sequence you wish to open. You will be prompted to open the file into the editor in one of three ways: Append the upcoming sequence to the currently displayed sequence, Insert the sequence at the cursor, or Replace the current sequence entirely.

After importing the sequence, click either the **Save** or **Exit** button. You are prompted to save the file under the current sequence name, the imported file name, or to display without saving. Once you make the save decision, you are prompted to import another sequence or return to the Sequence Editor.

10.3.1.3 Save

Use **Save** to save a copy of the sequence in a file on your hard drive.

All sequences introduced via the Open button become part of the current sequence. If the current sequence does not have a file name (i.e., it is a new sequence) you will be prompted to give that file a name upon activating the Save or Exit buttons. Otherwise, the contents of the editor will be saved to the current file without prompting.

CAUTION: The Sequence Editor is always in the Read/Write mode. Any keystrokes may alter the sequence!

10.3.1.4 Save As

Save As allows you to save the contents of the Sequence Window with a different name. You can then use the file in a different application or for another purpose. When you select the **Save As** command, the program prompts you to enter the new file name you would like to give the sequence. PatentIn v. 2.0 automatically adds the .GBS file extension. If you type the name of a file that already exists, the program prompts you to change the name of the new file.

10.3.1.5 Cancel

Cancel allows you to close the **FILE** menu options.

10.3.2 Attributes

The **Attributes** button calls the same sequence modification buttons as found on the Sequence Data browse screen and described in the paragraphs: **9.1.1 Organism**, **9.1.2 Feature**, and **9.1.3 Publications**.

10.3.3 SEARCH Menu

The commands in this menu allow you to locate a specific series of residues within your sequence. The menu options are **Find**, **Next**, and **Cancel**.

10.3.3.1 Find

The **Find** option opens a dialog box in which you enter the residue to find. The residue you want to find is called a **Search String**. **Find** command searches the sequence from the first position forward to the first occurrence of the Search String. You may type up to 10 characters.

Click OK once you have finished typing the characters. PatentIn v. 2.0 searches through the sequence looking for the characters you typed, ignoring spaces or line breaks. If the program finds the series, the cursor moves to the first residue of the series. If the program does not find the series, the cursor remains where it was before you selected the **Find** command and a message reports that the string was not found.

10.3.3.2 Next

The **Next** option command searches for subsequent occurrences of the set of residues you found last. Using **Next** eliminates the need to re-type the set of residues each time you want to find them.

10.3.3.3 Cancel

Cancel allows you to close the **SEARCH** menu options.

10.3.3.4 Nucleotide Codes Table

PatentIn v. 2.0 uses the following nucleotide symbols:

<i>Code</i>	<i>Group</i>	<i>Nucleotide(s)</i>
A	a	Adenine
C	c	Cytosine
G	g	Guanine
T	t	Thymine (in DNA)
U	u	Uracil (in RNA)
Y	c or t(u)	Pyrimidine
R	a or g	Purine
M	a or c	Amino
K	g or t(u)	Keto
S	g or c	strong interaction (3 hydrogen bonds)
W	a or t(u)	weak interaction (2 hydrogen bonds)
H	a or c or t(u)	not-G
B	g or t(u) or c	not-A

V	g or c or a	not-T or not-U
D	g or a or t(u)	not-C
N	g,a,c or t(u)	Any

10.3.3.5 Amino Acid Codes Table

PatentIn v. 2.0 uses one-letter amino acid in its Sequence Editor. They will appear in the printed sequence listing for features that require three letter amino acids.

NOTE: The three-letter code is shown here for reference. It is not used in the PatentIn v. 2.0 Sequence Editor at any time.

One-Letter Code	Three-Letter Code	Amino Acid
A	Ala	Alanine
R	Arg	Arginine
N	Asn	Asparagine
D	Asp	Aspartic acid
C	Cys	Cysteine
Q	Gln	Glutamine
E	Glu	Glutamic acid
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
L	Leu	Leucine
K	Lys	Lysine
M	Met	Methionine
F	Phe	Phenylalanine
P	Pro	Proline
S	Ser	Serine
T	Thr	Threonine
W	Trp	Tryptophan
Y	Tyr	Tyrosine
V	Val	Valine
B	Asx	Aspartic acid or Asparagine
Z	Glx	Glutamic acid or Glutamine
X	Xaa	Any amino acid

10.4 EXITING

When you have finished entering the sequence information, click either the **Save** or **Exit** buttons. If you click on the **Save** button, the information entered is saved to the current application. If you click on the **Exit** button, the program will automatically save the entered information, then return to the Sequence Data browse screen.

10.5 REORDERING SEQUENCES

PatentIn offers an easy way to re-arrange the order of sequences within the Sequence Listing by using the standard “Drag and Drop” mouse function.

Within the *Sequence Data Browse Screen*, left click on the sequence that you wish to move. While holding the left mouse button down, drag the sequence to the position within the list where you wish to place the sequence. You will notice an icon of a hand moving the sequence to its new location.

NOTE: The numerical order of the sequences within the Sequence Data Browse Screen denotes the order in which sequences are compiled in the Sequence Listing. The order number does not necessarily denote the actual SEQ ID NO order. If a sequence had previously been deleted (using the delete button on the screen), or a sequence has a CDS feature, which will in turn generate a protein sequence, the SEQ ID NO assigned to a sequence in a compiled Sequence Listing will not correspond to the order number shown in the list.

10.6 CODE ‘000’

10.6.1 Purpose

WIPO Standard ST.25, paragraph 5, states “Each sequence shall be assigned a separate sequence identifier. The sequence identifiers shall begin with 1 and increase sequentially by integers. If no sequence is present for a sequence identifier, the code ‘000’ should appear under the numeric identifier <400>, beginning on the next line following the SEQ ID NO.”

The purpose of this clause is to permit flexibility in the preparation and amendments of Sequence Listings. The characters 000 are the language-neutral code meaning “This sequence is intentionally skipped.”

In rare instances, this function may need to be used if, for instance, the sequence numbering within the Sequence Listing or patent application’s specification is modified, and a SEQ ID NO is dropped. In lieu of retyping the entire specification to accommodate a change in the SEQ ID NO order, the “000” function can be used as a placeholder within the Sequence Listing to preserve the current numbering.

The following example shows how sequence number 5 is intentionally skipped:

```
<210> 5  
<400> 5  
000
```

*NOTE: This differs from the “Delete” sequence button on the Sequence Data Browse Screen. (See Section 9.1.10 “Delete” for information on this button). When the “Generate Sequence Listing” button is used, the sequences that have been “deleted” by using the delete button on the Sequence Data Browse Screen will never show up on the compiled Sequence Listing. On the contrary, the “000” function is creating a **place***

holder for preserving a SEQ ID NO, and this will display on the compiled Sequence Listing.

10.6.2 How to insert the “000” code

If a sequence has already been included in the Sequence Listing, and it is intended for that SEQ ID NO to display the “000” code, first select the MODIFY button on the *Sequence Data Browse Screen*. Insert as the first three characters the code **000**. Then SAVE and EXIT. See example below.

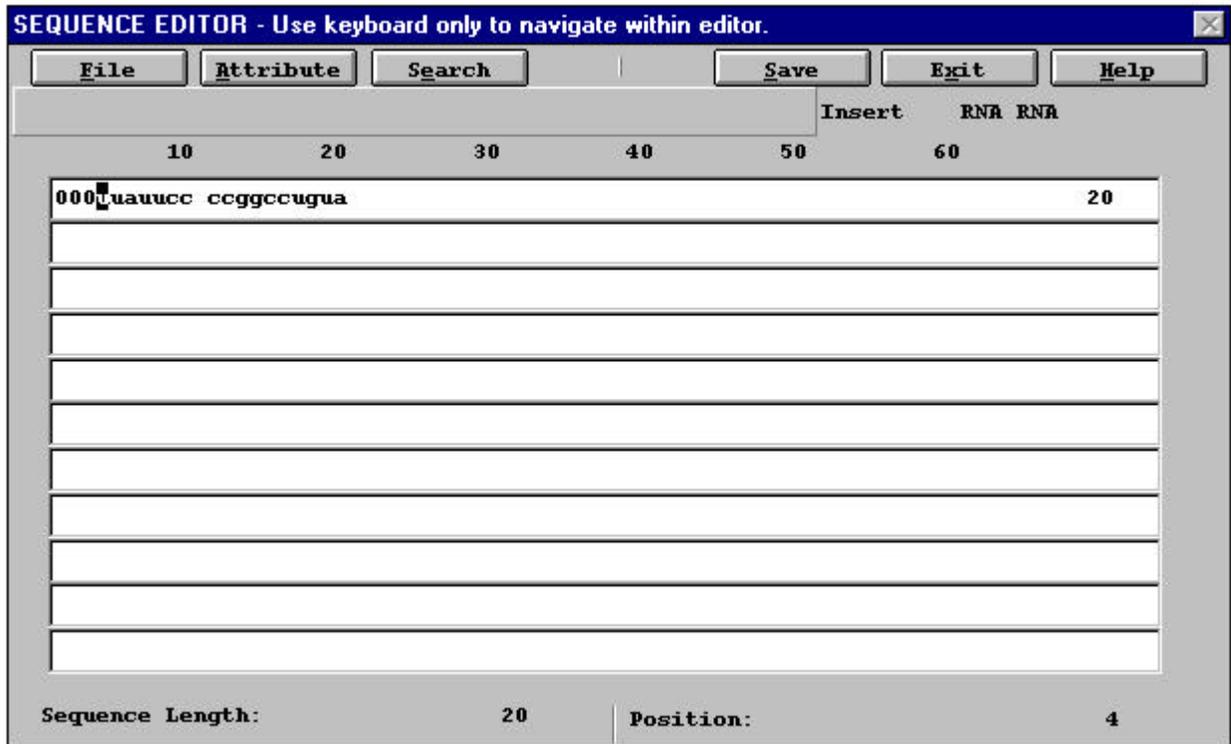


Figure 10-2: Example of 000 Place Holder

If a sequence has not already been entered, and a “000” placeholder needs to be inserted, first click on ADD within the *Sequence Data Browse Screen*. Then simply type **000** in place of the sequence. Then SAVE and EXIT.

If the “000” placeholder ever needs to be removed, click on the Modify button within the *Sequence Data Browse Screen*. Then delete the first three **000** characters. Then SAVE and EXIT. When you are returned to the Sequence Data Browse Screen, highlight the empty sequence and click on the Delete button.

11. Completing the Feature Information

The Feature Information data entry screen is a vital part of PatentIn v. 2.0. It has been designed to enter simple features easily. The data entry screen allows you to indicate biologically significant regions of the sequence. Annotating the features of your sequence enhances the utility of your data for other researchers.

The feature table is based exclusively based on Tables 5 and 6 within WIPO ST 25. Any feature can be represented, including, but not limited to, regions which:

- are promoters,
- are enhancers,
- are an enzyme's catalytic site,
- bind DNA,
- bind protein,
- result from recombination of different sequences,
- are a recognizable repeated unit,
- have secondary or tertiary structure,
- exhibit variation, or
- have been revised or corrected.

11.1 FEATURE INFORMATION DATA ENTRY SCREEN

Click on the **Feature** button in the Sequence Data browse screen to access the Feature Information data entry screen. The organization and interaction of the Feature Information data entry screen differs from the other data entry screens in the program. The Feature Information data entry screen is divided into three parts, as shown in **Figure 11-1**. These three parts are: **Data Entry Block**, **Other Information Window**, and **Feature Location Window**.

Features for SEQ1.GBS

Clear All Save Exit Help

Feature Name/Key-> modified_base Add'l Information...

Location: From-> 1 To-> 2 Complementary Strand->

Other Information:

cmm5s2u

Edit Feature Delete

Feature	Location
a CDS	(1)..(45)
b -35_signal	(2)..(10)
c -35_signal	(1)..(3)
d -10_signal Complement	((1)..(2))
e modified_base (1)..(2)	[] cmm5s2u
f - ()	[] Description of Unknown Organism:Unknown bacterium

Figure 11-1: Feature Information Data Entry Screen

11.1.1 Data Entry Block

The Data Entry Block consists of specific fields located at the top of the screen that display information about the feature that you are entering. Use the Data Entry Block of the Feature Information data entry screen to enter the type of feature and its location in the sequence. The Data Entry Block includes the following four fields: a **Feature Name/Key** field, a **From** field, a **To** field, and a **Complement** field.

11.1.1.1 Feature Name/Key

In this field, choose a keyword describing the feature from the pulldown list. This pulldown list will differ depending upon the type of sequence (i.e. nucleic acid versus protein). For more information on selecting the Feature Name/Key, refer to **Section 11.2** "Choosing a Feature Name/Key."

11.1.1.2 From Field

The next field is the **From** field in which you must enter the starting position of the feature.

11.1.1.3 To Field

The next field is the **To** field in which you must enter the ending position of the feature.

11.1.1.4 Complement Field

This field is used to indicate whether the feature is on the complementary strand. If so, select “C” from the pulldown list. Otherwise, leave blank or (if “C” has been selected) select the blank entry from the list. When “C” is selected, the word “**Complement**” appears in the Feature Location entry when saved.

***NOTE:** You should still specify locations on the complementary strand in the increasing 5’-to 3’ direction as if they were on the presented strand. That is, the number of the **From** position should always be less than that of the **To** position.*

11.1.2 Other Information Window

The middle third of the screen contains a window in which you may enter a more detailed description of the feature key. Use this **Other Information** window to enter specific feature qualifiers.

11.1.3 Feature Location Window

This window in the bottom third of the screen is used to specify the feature's location. For more information on the feature location, refer to **Section 11.3** “Specifying the Feature’s Location.”

Figure 11-1 illustrates how features are displayed in the Feature Location Window after they are saved. Special notation markers indicate a complex location in your feature. The word **Complement**, placed in front of any nucleic acid feature in which a **C** was entered in the field, indicates that the feature is on the complementary strand.

11.2 CHOOSING A FEATURE NAME/KEY

The data banks have developed a set of keywords that describe the major features found in nucleic acid and protein sequences. These keywords are referred to as feature keys. Feature keys indicate the biological nature of the annotated feature or other versions of the sequence. Choosing the appropriate feature key is important because these keys are frequently used to find or retrieve similar features or features with related functions.

PatentIn v. 2.0 provides lists of these feature keys; you must choose the one that best describes the feature you are entering. The lists of features differ depending upon whether your sequence is a nucleic acid or a protein. PatentIn v. 2.0 automatically provides the list based on the sequence type declared when the sequence was **Added** or **Imported** to PatentIn. If **DNA**, **RNA**, or **Combined DNA/RNA** was used, PatentIn v. 2.0 allows you to enter nucleic acid features; if a **PRT code** was used, you will be able to enter protein features.

11.2.1 Specific Feature Keys

With the cursor in the **Feature Name/Key** field, click in the field to see a list of feature keys.

Choose **misc_feature** if you cannot find a key that matches the feature you want to enter. Further define the **misc_feature** in the “**Other Information**” window. In the event that

you wish to only insert text (free text) in the “**Other Information**” window, select the **dash (-)** feature. Selecting **place holder (-)** automatically enters “N/A” in both the **From** and **To** fields.

See **Sections 13.1** and **13.2** for listings of other valid Feature Name keys.

11.2.2 Additional Information Button

If the selection in the **Feature Name/Key** field is **Mod_Base** for a nucleic acid sequence, or **Mod_Res** or **LIPID** for an amino acid sequence, pick lists are presented that permit additional features to be defined for the selection as per WIPO ST. 25. The results of this secondary selection appear in the Additional Information window.

When one of these features requiring additional features is selected, the Additional Information button is also enabled. This Additional Information button allows you to reenter the associated additional feature picklist to change your entry. See Feature Key Additional Descriptions for a list of the selection for each of the three features.

11.3 SPECIFYING THE FEATURE'S LOCATION

Once you have selected a feature key, you must indicate where in the sequence the feature occurs. PatentIn version 2.0 permits you to enter features that have single, distinct start and end points.

The feature location can be either a single residue or a contiguous span of residues.

You can enter a location by typing the position of the first residue in the feature in the **From** field and the position of the last residue in the feature in the **To** field.

Some features have a single residue for their location. If this is the case, enter the single position in both the **From** field and **To** field. PatentIn will enter “N/A” in the **To** field once the feature is saved.

Complex locations, such as “residues 9, 247, and 893, comprise the catalytic site,” may be entered by selecting “-” (dash) within the **Feature Name/Key** field and typing the description of the complex feature within the “**Other Information**” field <223>.

11.4 SAVING THE FEATURE

After you have entered the **Feature Key**, **Feature Location**, and **Other Information**, click either the **Save** or **Exit** button. If you click on the **Save** button, the information entered is saved to the current application and the fields are cleared for the next feature. If you click on the **Exit** button, the program will prompt you to save the entered information, then returns you to the Sequence Data browse screen. If you wish to interrupt the process of entering a feature without exiting the screen, activating **Clear All** will clear all of the data entry fields.

11.5 RECEIVING AN ERROR OR WARNING MESSAGE WHEN SAVING - THE NEXT ACTION TO TAKE

If your sequence contains a CDS or EXON that cannot be evenly grouped into codons (sets of three), PatentIn v. 2.0 displays an error message reminding you that “The final nucleotide in coding region lies within a codon” and instructs you to “revise the <value> in the **To** field.”

11.6 EDITING A FEATURE

If you wish to change a feature you have already entered, move the cursor to the **Feature Location Window**. Use the arrow keys to move up or down within the list of features and position the cursor on the tag assigned to the feature you wish to edit, then press **Enter** or activate the **Edit Feature button**. You may also double-click on the record to recall. The tag is found in the far-left column in the Feature Location Window. These alphabetical characters are permanently assigned sequentially by PatentIn from “a” to “zz” and do not change if a feature is deleted. Thus, the maximum number of features per sequence is 649.

The information pertaining to that feature appears in the Data Entry Block and in the **Other Information Window**. Make the necessary changes, then save the feature according to the instructions provided in **Section 11.4** “Saving the Feature.”

11.7 DELETING A FEATURE

Before deleting a feature you have already entered, be sure to complete and save the feature you are currently editing. The Data Entry Block should be empty before deleting features.

If you wish to delete a feature that you have already entered, highlight the feature in the **Feature Location Window** and either click the **Delete** button or press the **Delete** key on the keyboard. The program asks you to confirm that you really want to delete this feature. If you answer **Yes**, the program deletes the feature. If you select **No**, the program does not delete the feature.

11.8 EXITING THE FEATURE INFORMATION DATA ENTRY SCREEN

To exit the Feature Information data entry screen, you must first do one of the following: Either save all features by clicking on the **Save** button, or clear all fields by clicking on the **Clear All** button, which will empty the Data Entry Block fields. Then **Exit** normally.

11.9 ANNOTATING A PROTEIN CODING REGION

If your sequence contains a protein coding region (CDS), you should provide as much information about the features associated with that CDS as possible. It is recommended that you add detailed comments in the **Other Information Window**.

Describe a CDS on the Feature Information data entry screen by entering the features associated with it in the following order:

1. Select the feature "CDS" from the Feature Name/Key list.
2. Enter the start and end position for the region to be translated in the **FROM** and **TO** fields.
3. Enter any additional information in the **Other Information** window.
4. Save.
5. If needed, a "mat_peptide" feature can be added to display where the mature peptide starts in the sequence. Enter the starting position in both the **FROM** and **TO** fields.
Save and **Exit**.

SEQUENCE LISTING

```

<110> Doe, John
<120> Title for sample Sequence Listing
<130> Docket No.123123
<140>
<141>

<160> 2

<170> PatentIn Ver. 2.0

<210> 1
<211> 154
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(102)

<220>
<221> mat_peptide
<222> (43)

<400> 1
atg gcg aac aag ctc ttc ctc gtc tgc gca gct tgg agc cct gtg ttt      48
Met Ala Asn Lys Leu Phe Leu Val Cys Ala Ala Trp Ser Pro Val Phe
                -10                -5                -1   1

cat cct cac cac cgc ttc cgt cta tcg cac cgt tgt cga gtt cca cga      96
His Pro His His Arg Phe Arg Leu Ser His Arg Cys Arg Val Pro Arg
                5                10                15

aga ttg agacgccact aaccccatag gcccaataca gaaatgtcag aaggagtttt ca 154
Arg Leu
                20
    
```

```

<210> 2
<211> 34
<212> PRT
<213> Homo sapiens

<400> 2
Met Ala Asn Lys Leu Phe Leu Val Cys Ala Ala Trp Ser Pro Val Phe
          -10                -5                -1  1

His Pro His His Arg Phe Arg Leu Ser His Arg Cys Arg Val Pro Arg
          5                10                15

Arg Leu
      20
    
```

11.10 ANNOTATING AN INTERNAL STOP CODON IN A PROTEIN CODING REGION

If a nucleic acid sequence contains a protein coding region (CDS), the program will translate the CDS into its amino acid sequence in the Sequence Listing. In the rare case that a CDS contains an internal stop codon, you should follow the procedure below to annotate the region properly. This procedure is necessary for the program to accurately determine the numbering and length of the accompanying amino acid sequence.

Include the internal stop codon in the location of the CDS. When the program encounters an internal stop codon during translation, it places a blank space below it and includes the internal stop codon in the numbering and length calculation of the amino acid sequence (in this case, 33 amino acids in length). The terminal stop codon is excluded from the translation, numbering, and length of the amino acid sequence.

For example:

```

ATG GCG AAC AAG CTC TTC CTC GTC TGC GCA GCT TGA GCC CTG TGT TTC   48
Met Ala Asn Lys Leu Phe Leu Val Cys Ala Ala   Ala Leu Cys Phe
  1                5                10                15

ATC CTC ACC ACC GCT TCC GTC TAT CGC ACC GTT GTC GAG TTC CAC GAA   96
Ile Leu Thr Asn Ala Ser Val Tyr Arg Thr Val Val Glu Phe Asp Glu
          20                25                30

GAT TGA GACGCCAGTA ACCCCATAGG CCCAATACAG AAATGTCAGA AGGAGTTTCA   152
Asp
    
```

11.11 EXAMPLE OF SEQUENCE ANNOTATION

PatentIn v. 2.0 sequence file.

In the example above, PatentIn V. 2.0 will create 2 protein sequences below this DNA sequence. The first protein will include amino acid proteins 1 through 11. (**Met** through **Ala**). The second amino acid sequence that is generated will include positions 13 through 33 (**Ala** to **Asp**).

12. Generate an Application Sequence Listing

12.1 GENERATING THE SEQUENCE LISTING

When you have entered your sequence, along with all the relevant annotation and biological information, you can prepare the Sequence Listing that will be submitted to the receiving office. Follow the steps below to generate the Sequence Listing.

1. Select Generate Sequence Listing from the *Sequence Functions* submenu, or select “Generate Sequence Listing” from the *Quick Guide*.
2. PatentIn v. 2.0 verifies that the values have been entered in the general information-related fields that are mandatory for the Sequence Listing. If any of the mandatory information is missing, the program warns you that there are missing items and displays a list of missing items. (Refer to **Section 6.2** “Mandatory Fields in Data Entry Screens”).
3. If there are errors return to the appropriate screen(s) or the Sequence Editor and supply the missing data. Recreate the sequence listing by repeating steps 1 and 2.
4. After you have filled in all the mandatory fields, PatentIn v. 2.0 begins to compile the Application Sequence Listing and displays the message: “**Gathering Application Information, please wait...**”

If your sequence contains a CDS with an internal stop codon, the program displays a warning that it found an internal stop codon in the coding region, and then continues with the translation. If your CDS should NOT contain an internal stop codon, you may need to return to the Sequence Editor to check that the sequence is correct, or to the Feature Information data entry screen to check that the CDS was entered correctly. If your CDS actually does include an internal stop codon, select the **OK** button within the pop-up warning messages, and the system will proceed in building the Sequence Listing.

***NOTE.** If you make any changes to your application, you will need to recreate the Sequence Listing from the beginning by selecting **Generate Sequence Listing** from the Quick Guide.*

12.2 BROWSE SEQUENCE LISTING

Once the sequence is compiled after selecting the **Generate Sequence Listing** option, the complete Sequence Listing file will appear. From the “View Listing” window, the user may **Print, Copy to Disk, or Exit**. (If error messages are displayed at the top of the Sequence Listing, the user will be unable to copy the file to a disk.)

If a file has previously been through the **Generate Sequence Listing** function, and the Sequence Listing is error-free, the last selection on the *Quick Guide*, **Browse, Print, or Copy Sequence Listing** can be used. The purpose of having a separate option for browsing, printing, and copying an error free sequence listing is to avoid generating the Sequence Listing again. That can be a long process for a big file.

12.3 PRINTING THE SEQUENCE LISTING

PatentIn v. 2.0 stores the Application Sequence Listing in a file called *FILENAME.APP*. This is an ASCII file without page breaks so that the file can be printed on the widest variety of printers. You can print this file as the paper version of the application. This option is available in two places: the **Generate Sequence Listing** option and the **Browse, Print, or Copy Sequence Listing** option from the Browse Sequence Listing. To print, click on the print button and make your selection in the print dialog box.

WIPO ST.25 paragraph 3 requires that you submit a paper copy of your .APP file as part of your patent application. To aid the user in complying with the paper copy requirements of the Sequence Listing, PatentIn provides a function that numbers the printed pages. Since some receiving offices do not allow numbered pages (e.g., JPO), PatentIn v. 2.0 gives the user the selection to print with or without printed page numbers.

To change the default paper size in your printer, adjust the settings in your Windows Control Panel.

12.4 EDITING THE SEQUENCE LISTING

After you have generated the Sequence Listing, you should print the file and review it for accuracy. If information is missing or incorrect, supply or correct the information, then re-compile the Sequence Listing via the **Generate Sequence Listing** function.

CAUTION: Do NOT open the .APP file into a word processor. All editing must be done in an ASCII editor. A word processor will change the font and margins settings and therefore change the format of the output. It is suggested that you use the DOS "Edit" feature. Directions: Open a MSDOS Shell Command (COMMAND.COM), and at the prompt, type in the **EDIT** and the path to the file.
(i.e. **EDIT C:\PATIN2\SAMPLE\ SAMPLE.APP.**)

After editing your file, select, **Save**, then **Exit** (within the file pulldown menu). This is the fastest, safest way to edit an ASCII file, and will not interfere with PatentIn's formatting.

Notepad is also universally available within Windows™. In Windows 3.x, if Notepad is not in a Windows group, execute it from within the File Manager (which is a standard item in the Main group). The NOTEPAD.EXE file is in the Windows subdirectory. In Windows 95 and NT, it is on the Task Bar under the accessories. Open the .APP file as with any other Windows application. Notepad, like DOS EDIT, will not affect the formatting of the file.

CAUTION: In Notepad, if the file is saved to another directory using the "Save As" feature, ensure that the file is saved with the "Save as type" feature set to "All Files" to avoid a .TXT file extension.

12.5 ABOUT SEQUENCE LISTING OUTPUT

If the Sequence Listing contains a nucleotide sequence, PatentIn v. 2.0 groups bases in sets of 10 with 60 bases per line in the sequence output, except where coding regions are indicated. In coding regions, PatentIn v. 2.0 groups bases by three into codons with 16 codons per line.

If the Sequence Listing contains a protein sequence, PatentIn v. 2.0 places 16 amino acids per line in the sequence output.

If you select a DNA or RNA sequence that contains protein coding regions (CDS regions) for the Sequence Listing, PatentIn v. 2.0 translates the CDS; the corresponding three-letter amino acid code appears below each codon in the sequence output. In addition, the program automatically appends the three-letter amino acid translation of each CDS to the Sequence Listing as a separate sequence immediately following the nucleic acid sequence having the CDS feature. If you select one sequence with three CDS regions, your Sequence Listing will contain four sequences: the original nucleic acid sequence you selected, and the translated amino acid sequence comprising the three CDS regions.

If the application contains an amino acid sequence, the one-letter code used by the Sequence Editor will be converted to three-letter amino acid code in the output.

12.5.1 Numbering Sequence Listing Residues in the Output

For Sequence Listings that contain nucleic acids, PatentIn Version 2.0 counts the number of bases in each line, from the first base to the last base in the sequence, and places this number in the right margin of the sequence output. For Sequence Listings that contain amino acid sequences, PatentIn Version 2.0 numbers every fifth residue, beginning with the first amino acid, and places this number below the residue in the sequence output.

If the nucleic acid sequence contains a CDS region, PatentIn Version 2.0 numbers the first amino acid of the CDS region as 1. If the sequence contains multiple non-overlapping CDS regions, the program translates and numbers each CDS region independently. Each CDS translation then appears as a separate sequence in the output. In the rare case of overlapping CDS regions, the program will include only the first overlapping CDS region listed in the feature table in the sequence output. You will need to enter and save the remaining overlapping CDS regions as separate sequences in the Sequence Editor and include each one in the Sequence Listing.

If the CDS region includes a `mat_peptide` feature, the program numbers the first amino acid of the **mat_peptide** as 1, with negative numbers to the left (amino-terminal direction), extending to the beginning of the CDS. If the CDS contains multiple **mat_peptides**, the program numbers the first amino acid of the 5'-mat **mat_peptide** as 1.

12.6 INFORMATION ON HOW DATA IS STORED

All of the patent annotation and feature information is stored in a file with the same name you assigned to the application, followed by the file name extension `.APP`. Sequence data is stored in files with the extension `.GBS`. PatentIn v. 2.0 uses these two files to create the

computer-readable output file that will be sent to the receiving office as the Application Sequence Listing.

12.7 FINAL STEPS

1. Copy the Sequence Listing to selected medium (diskette).
2. Print out Sequence Listing (preferably using page numbers), on appropriate sized paper. PCT requires A4 paper other receiving offices (e.g. US PTO) accept 8 ½" x 11" paper.
3. Submit diskette and paper copy of Sequence Listing to appropriate receiving office.

13. Reference

13.1 FEATURE KEYS - NUCLEIC ACID

To select a feature key (in the Feature Information data entry screen), activate the Feature Name/Key pulldown list. The features on the list are arranged alphabetically to facilitate access, but the following section lists each feature on this list by category in order to demonstrate their biological relationships.

Feature Key Categories

RNA, transcript, and gene regions

Table 5: List of Feature Keys Related to Nucleotide Sequences

<i>Key</i>	<i>Description</i>
Allele	differs from the presented sequence at this location (and perhaps others)
Attenuator	(1) region of DNA at which regulation of termination of transcription occurs, which controls the expression of some bacterial operons; (2) sequence segment located between the promoter and the first structural gene that causes partial termination of transcription
C_region	constant region of immunoglobulin light and heavy chains, and T-cell receptor alpha, beta, and gamma chains; includes one or more exons depending on the particular chain
CAAT_signal	CAAT box; part of a conserved sequence located about 75 bp up-stream of the start point of eukaryotic transcription units which may be involved in RNA polymerase binding; consensus=GG (C or T) CAATCT
CDS	coding sequence; sequence of nucleotides that corresponds with the sequence of amino acids in a protein (location includes stop codon); feature includes amino acid conceptual translation
Conflict	independent determinations of the "same" sequence differ at this site or region
D-loop	displacement loop; a region within mitochondrial DNA in which a short stretch of RNA is paired with one strand of DNA, displacing the original partner DNA strand in this region; also used to describe the displacement of a region of one strand of duplex DNA by a single stranded invader in the reaction catalyzed by RecA protein
D-segment	diversity segment of immunoglobulin heavy chain, and T-cell receptor beta chain
enhancer	a cis-acting sequence that increases the utilization of (some) eukaryotic promoters, and can function in either orientation and in any location (upstream or downstream) relative to the promoter
exon	region of genome that codes for portion of spliced mRNA; may contain 5'UTR, all CDSs, and 3'UTR
GC_signal	GC box; a conserved GC-rich region located upstream of the start point of eukaryotic transcription units which may occur in multiple copies or in either orientation; consensus=GGGCGG
Gene	region of biological interest identified as a gene and for which a name has been assigned
Idna	intervening DNA; DNA which is eliminated through any of several kinds of recombination
intron	a segment of DNA that is transcribed, but removed from within the transcript by splicing together the sequences (exons) on either side of it

J_segment	joining segment of immunoglobulin light and heavy chains, and T-cell receptor alpha, beta, and gamma chains
LTR	long terminal repeat, a sequence directly repeated at both ends of a defined sequence, of the sort typically found in retroviruses
mat_peptide	mature peptide or protein coding sequence; coding sequence for the mature or final peptide or protein product following post-translational modification; the location does not include the stop codon (unlike the corresponding CDS)
misc_binding	site in nucleic acid which covalently or non-covalently binds another moiety that cannot be described by any other Binding key (primer_bind or protein_bind)
misc_	feature sequence is different from that presented in the entry and cannot be
difference	described by any other Difference key (conflict, unsure, old_sequence, mutation, variation, allele, or modified_base)
misc_feature	region of biological interest which cannot be described by any other feature key; a new or rare feature
misc_recomb	site of any generalized, site-specific or replicative recombination event where there is a breakage and reunion of duplex DNA that cannot be described by other recombination keys (iDNA and virion) or qualifiers of source key (/insertion_seq, /transposon, /proviral)
misc_RNA	any transcript or RNA product that cannot be defined by other RNA keys (prim_transcript, precursor_RNA, mRNA, 5'clip, 3'clip, 5'UTR, 3'UTR, exon, CDS sig_peptide, transit_peptide, mat_peptide, intron, polyA_site, rRNA, tRNA, scRNA, and snRNA)
misc_signal	any region containing a signal controlling or altering gene function or expression that cannot be described by other Signal keys (promoter, CAAT_signal, TATA_signal, -35_signal, -10_signal, GC_signal, RBS, polyA_signal, enhancer, attenuator, terminator, and rep_origin)
misc_structure	any secondary or tertiary structure or conformation that cannot be described by other Structure keys (stem_loop and D-loop)
modified_base	the indicated nucleotide is a modified nucleotide and should be substituted for by the indicated molecule (given in the mod_base qualifier value)
mRNA	messenger RNA; includes 5' untranslated region (5'UTR), coding sequences (CDS, exon) and 3' untranslated region (3'UTR)
mutation	a related strain has an abrupt, inheritable change in the sequence at this location
N_region	extra nucleotides inserted between rearranged immunoglobulin segments
old_sequence	the presented sequence revises a previous version of the sequence at this location
polyA_signal	recognition region necessary for endonuclease cleavage of an RNA transcript that is followed by polyadenylation; consensus=AATAAA
polyA_site	site on an RNA transcript to which will be added adenine residues by post-transcriptional polyadenylation
precursor_RNA	any RNA species that is not yet the mature RNA product; may include 5' clipped region (5'clip), 5' untranslated region (5'UTR), coding sequences (CDS, exon), intervening sequences (intron), 3' untranslated region (3'UTR), and 3' clipped region (3'clip)
prim_	primary (initial, unprocessed) transcript; includes 5' clipped region (5'clip),
transcript	5' untranslated region (5'UTR), coding sequences (CDS, exon), intervening sequences (intron), 3' untranslated region (3'UTR), and 3' clipped region (3'clip)
primer_bind	non-covalent primer binding site for initiation of replication, transcription, or reverse transcription; includes site(s) for synthetic, for example, PCR primer elements
promoter	region on a DNA molecule involved in RNA polymerase binding to initiate transcription

protein_bind	non-covalent protein binding site on nucleic acid
RBS	ribosome binding site
repeat_region	region of genome containing repeating units
repeat_unit	single repeat element
rep_origin	origin of replication; starting site for duplication of nucleic acid to give two identical copies
rRNA	mature ribosomal RNA; the RNA component of the ribonucleoprotein particle (ribosome) which assembles amino acids into proteins
S_region	switch region of immunoglobulin heavy chains; involved in the rearrangement of heavy chain DNA leading to the expression of a different immunoglobulin class from the same B-cell
satellite	many tandem repeats (identical or related) of a short basic repeating unit; many have a base composition or other property different from the genome average that allows them to be separated from the bulk (main band) genomic DNA
scRNA	small cytoplasmic RNA; any one of several small cytoplasmic RNA molecules present in the cytoplasm and (sometimes) nucleus of a eukaryote
sig_peptide	signal peptide coding sequence; coding sequence for an N-terminal domain of a secreted protein; this domain is involved in attaching nascent polypeptide to the membrane; leader sequence
snRNA	small nuclear RNA; any one of many small RNA species confined to the nucleus; several of the snRNAs are involved in splicing or other RNA processing reactions
source	identifies the biological source of the specified span of the sequence; this key is mandatory; every entry will have, as a minimum, a single source key spanning the entire sequence; more than one source key per sequence is permissible
stem_loop	hairpin; a double-helical region formed by base-pairing between adjacent (inverted) complementary sequences in a single strand of RNA or DNA
STS	Sequence Tagged Site; short, single-copy DNA sequence that characterizes a mapping landmark on the genome and can be detected by PCR; a region of the genome can be mapped by determining the order of a series of STSs
TATA_signal	TATA box; Goldberg-Hogness box; a conserved AT-rich septamer found about 25 bp before the start point of each eukaryotic RNA polymerase II transcript unit which may be involved in positioning the enzyme for correct initiation; consensus=TATA(A or T)A(A or T)
terminator	sequence of DNA located either at the end of the transcript or adjacent to a promoter region that causes RNA polymerase to terminate transcription; may also be site of binding of repressor protein
transit_peptide	transit peptide coding sequence; coding sequence for an N-terminal domain of a nuclear-encoded organellar protein; this domain is involved in post-translational import of the protein into the organelle
tRNA	mature transfer RNA, a small RNA molecule (75-85 bases long) that mediates the translation of a nucleic acid sequence into an amino acid sequence unsure author is unsure of exact sequence in this region
V_region	variable region of immunoglobulin light and heavy chains, and T-cell receptor alpha, beta, and gamma chains; codes for the variable amino terminal portion; can be made up from V_segments, D_segments, N_regions, and J_segments
V_segment	variable segment of immunoglobulin light and heavy chains, and T-cell receptor alpha, beta, and gamma chains; codes for most of the variable region (V_region) and the last few amino acids of the leader peptide

variation	a related strain contains stable mutations from the same gene (for example, RFLPs, polymorphisms, etc.) which differ from the presented sequence at this location (and possibly others)
3'clip	3'-most region of a precursor transcript that is clipped off during processing
3'UTR	region at the 3' end of a mature transcript (following the stop codon) that is not translated into a protein
5'clip	5'-most region of a precursor transcript that is clipped off during processing
5'UTR	region at the 5' end of a mature transcript (preceding the initiation codon) that is not translated into a protein
-10_signal	pribnow box; a conserved region about 10 bp upstream of the start point of bacterial transcription units which may be involved in binding RNA polymerase; consensus=TAtAaT
-35_signal	a conserved hexamer about 35 bp upstream of the start point of bacterial transcription units; consensus=TTGACa [] or TGTTGACA []

13.2 FEATURE KEYS - AMINO ACID

Table 6: List of Feature Keys Related to Protein Sequences

<i>Key</i>	<i>Description</i>
CONFLICT	different papers report differing sequences
VARIANT	authors report that sequence variants exist
VARSPLIC	description of sequence variants produced by alternative splicing
MUTAGEN	site which has been experimentally altered
MOD_RES	post-translational modification of a residue
ACETYLATION	N-terminal or other
AMIDATION	generally at the C-terminal of a mature active peptide
BLOCKED	undetermined N- or C-terminal blocking group
FORMYLATION	of the N-terminal methionine
GAMMA-CARBOXYGLUTAMIC ACID HYDROXYLATION	of asparagine, aspartic acid, proline or lysine
METHYLATION	generally of lysine or arginine
PHOSPHORYLATION	of serine, threonine, tyrosine, aspartic acid or histidine
PYRROLIDONE CARBOXYLIC ACID	N-terminal glutamate which has formed an internal cyclic lactam
SULFATATION	generally of tyrosine
LIPID	covalent binding of a lipidic moiety
MYRISTATE	myristate group attached through an amide bond to the N-terminal glycine residue of the mature form of a protein or to an internal lysine residue
PALMITATE	palmitate group attached through a thioether bond to a cysteine residue or through an ester bond to a serine or threonine residue
FARNESYL	farnesyl group attached through a thioether bond to a cysteine residue
GERANYL-GERANYL	geranyl-geranyl group attached through a thioether bond to a cysteine residue
GPI-ANCHOR	glycosyl-phosphatidylinositol (GPI) group linked to the alpha-carboxyl group of the C-terminal residue of the mature form of a protein
N-ACYL DIGLYCERIDE	N-terminal cysteine of the mature form of a prokaryotic lipoprotein with

<i>Key</i>	<i>Description</i>
	an amide-linked fatty acid and a glyceryl group to which two fatty acids are linked by ester linkages
DISULFID	disulfide bond; the 'FROM' and 'TO' endpoints represent the two residues which are linked by an intra-chain disulfide bond; if the 'FROM' and 'TO' endpoints are identical, the disulfide bond is an interchain one and the description field indicates the nature of the cross-link
THIOLEST	thiolester bond; the 'FROM' and 'TO' endpoints represent the two residues which are linked by the thiolester bond
THIOETH	thioether bond; the 'FROM' and 'TO' endpoints represent the two residues which are linked by the thioether bond
CARBOHYD	glycosylation site; the nature of the carbohydrate (if known) is given in the description field
METAL	binding site for a metal ion; the description field indicates the nature of the metal
BINDING	binding site for any chemical group (co-enzyme, prosthetic group, etc.); the chemical nature of the group is given in the description field
SIGNAL	extent of a signal sequence (prepeptide)
TRANSIT	extent of a transit peptide (mitochondrial, chloroplastic, or for a microbody)
PROPEP	extent of a propeptide
CHAIN	extent of a polypeptide chain in the mature protein
PEPTIDE	extent of a released active peptide
DOMAIN	extent of a domain of interest on the sequence; the nature of that domain is given in the description field
CA_BIND	extent of a calcium-binding region
DNA_BIND	extent of a DNA-binding region
NP_BIND	extent of a nucleotide phosphate binding region; the nature of the nucleotide phosphate is indicated in the description field
TRANSMEM	extent of a transmembrane region
ZN_FING	extent of a zinc finger region
SIMILAR	extent of a similarity with another protein sequence; precise information, relative to that sequence is given in the description field
REPEAT	extent of an internal sequence repetition
HELIX	secondary structure: Helices, for example, Alpha-helix, 3(10) helix, or Pi-helix
STRAND	secondary structure: Beta-strand, for example, Hydrogen bonded beta-strand, or Residue in an isolated beta-bridge
TURN	secondary structure Turns, for example, H-bonded turn (3-turn, 4-turn or 5-turn)
ACT_SITE	amino acid(s) involved in the activity of an enzyme
SITE	any other interesting site on the sequence
INIT_MET	the sequence is known to start with an initiator methionine
NON_TER	the residue at an extremity of the sequence is not the terminal residue; if applied to position 1, this signifies that the first position is not the N-terminus of the complete molecule; if applied to the last position, it signifies that this position is not the C-terminus of the complete molecule; there is no description field for this key
NON_CONS	non consecutive residues; indicates that two residues in a sequence are not

<i>Key</i>	<i>Description</i>
	consecutive and that there are a number of unsequenced residues between them
UNSURE	uncertainties in the sequence; used to describe region(s) of a sequence for which the authors are unsure about the sequence assignment

13.3 FEATURE KEYS - ADDITIONAL DESCRIPTIONS

Any information required by WIPO ST. 25 or any information that the user may wish to include clarifying a feature should be typed into the Additional Information Window before storing the feature.

For **mod_res**, **modified_base**, and **lipid** features the following additional qualifiers are presented as selection lists. The selections may be changed by activating the **Add'l Information** button, which is displayed with these features.

Feature: MOD_RES (modified residue)
ACETYLATION
AMIDATION
BLOCKED
FORMYLATION
GAMMA_CARBOXYGLUTAMIC ACID
METHYLATION
PHOSPHORYLATION
PYRROLIDONE CARBOXYLIC ACID
SULFATATION

Feature: LIPID
PALMITATE
FARNESYL
GERANYL-GERANYL
GPI-ANCHOR
N_ACYL DIGLYCERIDE

Feature: Modified_base	
Abbrev.	Full description
ac4c	4-acetylcytidine
chm5u	5-(carboxyhydroxymethyl)uridine
cm	2'-O-methylcytidine
Cmm5s2u	5-carboxymethylaminomethyl-2-thiouridine
cmm5u	5-carboxymethylaminomethyluridine
d	dihydrouridine
fm	2'-O-methylpseudouridine
gal q	beta, D-galactosylqueosine
gm	2'-O-methylguanosine
i	inosine
i6a	N6-isopentenyladenosine
m1a	1-methyladenosine
m1f	1-methylpseudouridine

m1g	1-methylguanosine
m1i	1-methylinosine
m22g	2,2-dimethylguanosine
m2a	2-methyladenosine
m2g	2-methylguanosine
m3c	3-methylcytidine
m5c	5-methylcytidine
m6a	N6-methyladenosine
m7g	7-methylguanosine
mam5u	5-methylaminomethyluridine
mam5s2u	5-methoxyaminomethyl-2-thiouridine
man q	beta, D-mannosylqueosine
mcm5s2u	5-methoxycarbonylmethyl-2-thiouridine
Mcm5u	5-methoxycarbonylmethyluridine
mo5u	5-methoxyuridine
ms2i6a	2-methylthio-N6-isopentenyladenosine
ms2t6a	N-((9-beta-D-ribofuranosyl-2-MePu-6-yl)carbamoyl)threonine
mt6a	N-((9-B-D-ribofuranosylPu-6-yl)N-methylcarbamoyl)threonine
mv	uridine-5-oxyacetic acid-methylester
o5u	uridine-5-oxyacetic acid (v)
osyw	wybutoxosine
p	pseudouridine
Q	queosine
S2c	2-thiocytidine
s2t	5-methyl-2-thiouridine
s2u	2-thiouridine
s4u	4-thiouridine
t	5-methyluridine
t6a	N-((9-beta-D-ribofuranosylpurine-6-yl)carbamoyl)threonine
tm	2'-O-methyl-5-methyluridine
um	2'-O-methyluridine
yw	wybutosine
x	3-(3-amino-3-carboxy-propyl)uridine, (acp3)u

13.4 LIST OF NUMERIC HEADING DEFINITIONS

- <110> Applicant
- <120> Title of invention
- <130> File reference
- <140> Current patent application
- <141> Current filing data
- <150> Earlier patent application
- <151> Earlier application filing date
- <160> Number of SEQ ID NOs

- <170> Software
- <210> Information for SEQ ID NO:# (show sequence # only)
- <211> Length
- <212> Type
- <213> Organism
- <220> Feature (main heading)
- <221> Name/Key
- <222> Location
- <223> Other Information
- <300> Publication Information (main heading)
- <301> Authors
- <302> Title
- <303> Journal
- <304> Volume
- <305> Issue
- <306> Pages
- <307> Date
- <308> Database name/accession number
- <309> Database entry date
- <310> Document number
- <311> Filing date
- <312> Publication date
- <313> Relevant residues FROM .. TO ..
- <400> Sequence (sequence shown under line after sequence #)

000 Annotation	39	Filing Date.....	20
Add.....	17, 18, 21, 22, 23, 26, 29, 30, 31, 33	First Name.....	17
Add'l Information Button	44, 58	From	28, 42
Amino Acid Codes	38	Issue	28
Applicant Information	17	Last Name	17
Applicant Information Browse Screen	18	Mandatory	14
Application Number	20	Middle Initial.....	18
Artificial Sequence.....	25	Publication Date	28
Author Information	29	Publication Name	28
Browse Screen		Publication Type.....	27
Applicant Information	18	Relevant Residues.....	28
Prior Application Information.....	22	Start Page	28
Clear All Button.....	45	Suffix	17
Complement Field.....	43	Title.....	20, 28
Create New Application	9	To.....	29, 42
Current Application Information Menu Option .	6	Volume.....	28
Dash Feature	44	File Reference Field.....	9, 20
Data Entry Block.....	42	Filing Date Field.....	20
Data Entry Screen	6	Filters.....	32, 36
Applicant Information	17	DNA.....	32, 35
Application	6	PRT.....	32
Author Information.....	29	RNA.....	32, 35
Current Application Information.....	19	Find.....	<i>See</i> Sequence Editor, Search Menu
Cursor Control Keys.....	7	From Field.....	28, 42
Feature Information.....	25, 41, 46	Generate Sequence Listing.....	11, 39, 49, 50
Help.....	16	Help	16
Maneuvering	7	Import	23, 32
Prior Application Information.....	20	Individual	17
Sequence Functions.....	7	Installation	3
Delete	17, 18, 21, 22, 23, 29, 31, 33	Internal Stop Codon.....	47
DNA Filter.....	32, 35	Issue Field.....	28
End Page Field.....	28	Language Preference	9
Error Message.....	45	Languages	1
Exit.....	17, 18, 20, 22	Library of Previously Defined Publications.....	30
Feature7, 8, 23, 25, 33, 37, 41, 42, 43, 44, 45, 46, 49, 53, 56, 58		Limitations	8
Feature Button	41	List of Feature Keys Related to Nucleotide Sequences.....	53
Feature Information.....	41, 45	List of Feature Keys Related to Protein Sequences.....	56
Data Entry Screen.....	25	Lists	
Feature Information Data Entry Screen		Cursor Control Keys	7
Data Entry Line.....	42	Restricted Vocabulary	14
Feature Location Window.....	43	Main Menu.....	6
Other Information Window.....	43	Main Menu/Quick Guide	11, 13
Feature Key Categories	53	Mandatory Fields.....	14
Feature Key Descriptions	44, 56, 58	Mat_peptide	51
Feature Keys		Menu	
Nucleic Acid	53	Exit	12
Feature Name/Key Field.....	42, 43	Help	16
Field		Main.....	6, 11
Application Number	20	Main/Quick Guide	11
Complement.....	43	Modify	17, 18, 21, 22, 23, 26, 29, 30, 31, 33
Completing	14	New	<i>See</i> Sequence Editor, File Menu
End Page.....	28	New Application	9
Feature Name/Key.....	42, 43	Next	<i>See</i> Sequence Editor, Search Menu
File Reference	9, 20		

Next Button	26	Search String <i>See</i> Sequence Editor, Search Menu	
Nucleotide Codes	37	Sequence Annotation	47
Numbering Sequence Listing Residues in the		Sequence Data Browse Screen	23
Output	51	Sequence Editor.....	23, 29, 34, 51
Open	<i>See</i> Sequence Editor, File Menu	Accessing	35
Order	<i>See</i> Sequence Editor	Data Entry Block	45
Organism	7, 14, 15, 23, 24, 25, 37	Feature Location Window	45
Other.....	25	File Menu	35
Sequence	25	Menu Bar	34
Unknown	25	Order.....	23
Organization	17	Other Information Window	45
Other Organism.....	25	Reordering.....	39
Prior Application Information.....	20	Search Menu.....	37
Prior Application Information Menu Option	6	Sequence Entry Window	34
Protein Coding Region (CDS)	45, 47, 51	Sequence Entry Window.....	34
PRT Filter	32, 35	Sequence Listing	49
PRT One-Letter Filter	32, 35	Browse	7, 50
PRT Three-Letter Filter	33, 35	Editing	50
Publication Date Field	28	Generate	7
Publication Name Field	28	Generate	49
Publication Type Field	27	Output	50
Publications	7, 14, 23, 25, 30, 31, 33, 37	Printing	50
Regions		Specific Feature Keys	43
<i>Gene</i>	53	Start Page Field	28
<i>RNA</i>	53	System Requirements	2
<i>Transcript</i>	53	Table 5	53
Relevant Residues	28	Table 6	56
Reordering	<i>See</i> Sequence Editor	Title Field.....	20, 28
Restore.....	23, 31, 33	To Field.....	28, 42
RNA Filter	32, 35	Unknown Organism.....	25
Save	<i>See</i> Sequence Editor, File Menu	Volume Field.....	28
Feature Information	44, 45	Warning Message	45
Save As.....	<i>See</i> Sequence Editor, File Menu	WIPO ST.25.....	1, 25, 50