INN Antibody Annotation Format

Prof. Andrew C.R. Martin, 15th March 2017

Updated: 23.08.17, 19.04.18, 31.05.18, 11.09.18, 20.02.19, 04.11.19, 08.11.19, 02.04.20

Purpose of this document

The INN publications provide information on the drugs for which names are assigned. In the case of small molecule drugs this is generally straightforward with a drawing of the molecule and a standard chemical description. SMILES strings might also be used.

Description of biologics is more difficult – the amino acid sequence is only a partial description since post-translational modifications (including glycosylation and disulfide bonding) also come into play and some therapeutic agents may include multiple protein chains.

Antibodies are a subset of biologics that present a particular set of problems. In addition to their actual sequence, it is considered valuable to know the germline source of these molecules, as well as they way in which they have been produced (e.g. murine, human, humanized, resurfaced, chimeric).

Previous descriptions of antibodies used by the INN have attempted to co-opt the type of nomenclature used for small-molecule drugs. For such complex molecules, including all the required information makes these descriptions difficult to read and difficult to verify or parse automatically by computer.

The new format is simpler to read and to parse and allows a huge amount of flexibility.

The new description format

The description format consists merely of simple 'keyword-value' tuples which can be displayed on separate lines. Each keyword is immediately followed by a colon and values are followed by a semi-colon.

```
Keyword: value;
```

This enables all the tuples to be concatenated for a more compact display. The requirement that colons and semi-colons are only ever used at these positions (i.e. the actual keywords and values never contain these characters) makes automatic parsing and format verification trivial. This representation is also trivially extensible – new keywords can be introduced as required without affecting the overall format.

The sequence itself is represented with 50 amino acids per line with a space after each 10 amino acids. The accumulated number of amino acids is shown at the end of each line. Thus, for a chain of more than 243 amino acids, these numbers will be 50, 100, 150, 200, 243. The sequences are introduced with a header of 'Heavy Chain:' or 'Light Chain:' and the sequence is terminated by two forward slashes beginning a line by themselves. This is done to make parsing of the sequence data more straightforward (and is the same indicator used in other sequence formats such as EMBL-ENA, Genbank and UniProtKB/SwissProt.

The keyword-value format also allows more unusual cases such as bispecific antibodies and fusion antibodies to be handled in a straightforward manner. Standard antibodies have a single type of 'light' chain and a single type of 'heavy' chain. In the case of bispecifics (which have two different heavy and/or light chains) and fusions (where a single chain Fv antibody is fused to the C-terminus of the heavy chain), there may be two different light and/or heavy chain sequences. In such cases an instance qualifier (a number in square brackets) is used between the keyword and the colon to indicate the chain to which the keywords refer.

```
Keyword[1]: value;
```

With the exception of the `Format:` record which must appear first, there is no fixed order in which keywords must appear, but the order described in the examples below is regarded as the standard order and should be used for antibodies of the types described.

Note that all **residue numbering used in the file is sequential**. Kabat, Chothia, Martin, IMGT, or Aho numbering are **not** used.

Format for a standard antibody (e.g. 10387)

`Format:`	This describes the overall nature of the molecule.
	It starts with an optional indicator of non-standard antibodies:
	`[conjugated|bispecific]`
	This is followed by the optional keyword indicating fusion proteins:
	`[fusion]`
	This is followed by the source of the antibody:
	`[chimeric|humanized|resurfaced|human|mouse|llama|canine|felinized]`
	In some cases, (including normal antibodies, fusions and bispecifics) there may be more than one source – these are separated by a forward slash. For example:
	`human/humanized`
	When this occurs for a normal antibody, the `LightSource:` and `HeavySource:` fields are used to explain which is which.
	This is then followed by the words
	`'monoclonal [scFv|Fab|VHFc] antibody'`
	`and may then be followed by qualifiers`
	`‘hexamer’`

	For example:
	`'chimeric monoclonal antibody'`
	`'conjugated chimeric monoclonal antibody'`
	`'humanized monoclonal antibody'`
	`'human monoclonal antibody hexamer'`
	`'bispecific humanized monoclonal antibody'`
	`'conjugated humanized monoclonal antibody'`
	`'fusion human monoclonal antibody'`
	`'bispecific fusion human monoclonal antibody'`
	`'bispecific fusion human/humanized monoclonal antibody'`
	`'conjugated human monoclonal antibody'`
	`'human monoclonal antibody'`
	
	In the case of fusions with non-antibody domains, this is then followed by the key words:
	`'fusion with'`
	and the species and name of the fusion protein. e.g.
	`'fusion with human iduronate-2-sulfatase'`

`ID:`	Used to provide an original identifier for the antibody (e.g. clone name, lab code, etc)
`Note:`	Note fields may be used throughout the entry and refer to the previous record.
`HeavySource:`	Optional field used for normal antibodies where the two chains are of different origins (e.g. human, humanized, chimeric, etc). Used with `LightSource:` If there are multiple instances (i.e. non-standard antibody formats), then the `Source[]:` record will be used instead.
`LightSource:`	Optional field used for normal antibodies where the two chains are of different origins (e.g. human, humanized, chimeric, etc). Used with `HeavySource:` If there are multiple instances (i.e. non-standard antibody formats), then the `Source[]:` record will be used instead.

`Type:`	This is a summary of the antibody type describing the nature of both heavy and light chains. It consists of two parts joined with a dash. The first part describes the heavy chain:
	`[IgG1|IgG2|IgG3|IgG4|Fv]`
	while the second part describes the light chain class:
	`[kappa|lambda]`
	For example:
	`IgG1-kappa`

	`IgG4-lambda`
	In some cases, the heavy chain constant regions can come from more than one source. In such cases, the first part can have more than one type separated by slashes. For example:
	`IgG2/IgG4-kappa`
	Similarly the light chain can have more than one source. In such cases, the second part can have more than one type separated by slashes. For example (e.g. 10829):
	`IgG1-lambda/kappa`
	In addition there is an optional third part for special format chains. Currently these may be:
	`[Crossmab|scFv|Fab|scFab|VHFc|XscFv]`
	`(XscFv represents a crossed-over scFv otherwise known as DART)`
	In the case of an scFv, the heavy chain class information is irrelevant so for the light chain part we have
	`Type[1]: Fv-[kappa|lambda] scFv;`
	and for the heavy chain part we have
	`Type[2]: Fv-heavy scFv;`
	(e.g. 10562)
	Note that a ‘dimer’ keyword may also follow the ‘scFv’ keyword. (e.g. 10652, 10664)
	`Type type may also be specified as ‘single-domain’ for single heavy domain (nanobody) antibodies. (e.g. 11137)`
	The type may also be specified as ‘`OTHER`’ for non-antibody chains that are fused with a non-antibody protein (e.g. 10664, 11116)

`Conjugate:`	This field is only used when the antibody is conjugated to a drug molecule and indicates the name of that conjugated drug. Details of the linker are not included.
`Antigen:`	The name of the antigen (including the common name species as the first word) followed by any pseudonyms separated by commas and finally by the gene name in parentheses. Greek characters are expanded to their name (e.g. α becomes alpha). In the case of antibodies where an antigen binds across two instances (i.e. the light and heavy chains have separate instance numbers), these are

	indicated. This serves to show how instances are paired to form antigen combining sites. e.g.
	`Antigen[1,2]: CD123;`
	`Antigen[3,4]: CD3;`

`Binding:`	A more general form of 'Antigen:' used for fused non-antibody domains (e.g. anticalins). The name of the bound target (including the common name species as the first word) followed by any pseudonyms separated by commas and finally by the gene name in parentheses. Greek characters are expanded to their name (e.g. α becomes alpha).
	Since the annotation is for antibodies, this will require and instance to be supplied since this will be part of a fusion. e.g.
	`Binding[2]: CD137;`
	`(See 11094)`

`HeavyChainClass``:`	The heavy chain class is given as:
	`[Gamma1|Gamma2|Gamma3|Gamma4|NONE]`
	and must match the heavy chain specification given in the `Type:` field.
	
	This record is optional when the Type: is an Fv or ScFv. Alternatively the `NONE` type may be used when the Type: field heavy chain is specified as Fv.
	
	Where the heavy chain is a composite of more than one type, separate records are given for the different domains with the domain(s) given in parentheses. If there is an instance qualifier, this appears after the domain indicator. For example:
	`HeavyChainClass(CH1): Gamma2;`
	`HeavyChainClass(CH2,CH3): Gamma4;`
	`HeavyChainClass(CH1)[1]: Gamma1;`

`HeavyChainLength:`	The total number of amino acids in the heavy chain. In the case of fusions, this label is applied to the N-terminal antibody heavy chain. It is the full length of the chain including any fusion proteins. This field is replaced by `ChainLength`: for scFv/scFab/XscFv
`ChainLength:`	Used instead of `HeavyChainLength:` and `LightChainLength:` for cases such as scFv/scFab/XscFv (DART) proteins where you can't assign chains as (predominantly) heavy or light. It is the full length of the chain including (perhaps) light and heavy domains and any fusion proteins. Instance qualifiers are used to list the included entities.

`LightChainClass:`	The light chain class is given as:
	`[kappa|lambda]`
	and must match the light chain specification given in the `Type:` field.
	Where the variable and constant domains come from different germlines (e.g. 10829), each domain is indicated in parentheses. Where an instance qualifier is also used, this appears after the domain indicator. For example:
	`LightChainClass(VL): lambda;`
	`LightChainClass(CL): kappa;`
	`LightChainClass(CL)[1]: lambda;`

`LightChainLength:`	The total number of amino acids in the light chain. In the case of fusions, this label is applied to the N-terminal antibody light chain. It is the full length of the chain including any fusion proteins. This field is replaced by `ChainLength`: for scFv/scFab/XscFv
`HeavyPotentialNGlycos:`	This is a space-separated list of asparagine positions in the heavy chain that are part of N-{P}-[ST]-{P} patterns characteristic of N-linked glycosylation sites. ([] indicates alternatives and {} indicates amino acids other than those listed). In the case of fusion proteins, this label is applied to the N-terminal antibody heavy chain.
`HeavyConfirmedNGlycos:`	This is a space-separated list of asparagine positions in the heavy chain that are part of N-{P}-[ST]-{P} patterns characteristic of N-linked glycosylation sites and have been confirmed by the submitters as being glycosylated. In the case of fusion proteins this label is applied to the N-terminal antibody heavy chain.
`LightPotentialNGlycos:`	This is a space-separated list of asparagine positions in the light chain that are part of N-{P}-[ST]-{P} patterns characteristic of N-linked glycosylation sites. In the case of fusion proteins, this label is applied to the N-terminal antibody light chain.
`LightConfirmedNGlycos:`	This is a space-separated list of asparagine positions in the light chain that are part of N-{P}-[ST]-{P} patterns characteristic of N-linked glycosylation sites and have been confirmed by the submitters as being glycosylated. In the case of fusion proteins, this label is applied to the N-terminal antibody light chain.
HeavyPotentialOGlycos:	This is a space-separated list of possible O-linked glycosylation sites identified by the applicant.
HeavyConfirmedOGlycos:	This is a space-separated list of O-linked

	glycosylation sites confirmed by the applicant.
HeavyConfirmedPTM:	This is followed by the type of PTM (deamidation, oxidation, clipping, Nter-cyclization, etc.) and a space-separated list of positions in the heavy chain that have been confirmed as PTM sites. e.g. HeavyConfirmedPTM: deamidation 385 435; HeavyConfirmedPTM: oxidation 385; HeavyConfirmedPTM: cterclip 448; HeavyConfirmedPTM: nterpca 1;
LightConfirmedPTM:	This is followed by the type of PTM (deamidation, oxidation) and a space-separated list of positions in the light chain that have been confirmed as PTM sites.
ConfirmedPTM[x]:	This is followed by the type of PTM (deamidation, oxidation) and a space-separated list of positions in the light chain that have been confirmed as PTM sites.
PotentialNGlycos[x]:	This is a space-separated list of asparagine positions in a non-antibody chain that are part of N-`{P}`-`[ST]`-`{P}` patterns characteristic of N-linked glycosylation sites.
ConfirmedNGlycos[x]:	This is a space-separated list of asparagine positions in a non-antibody chain that are part of N-`{P}`-`[ST]`-`{P}` patterns characteristic of N-linked glycosylation sites and have been confirmed by the submitters as being glycosylated.
PotentialOGlycos[x]:	This is a space-separated list of possible O-linked glycosylation sites identified by the applicant.
ConfirmedOGlycos[x]:	This is a space-separated list of O-linked glycosylation sites confirmed by the applicant.
`HeavyCysPositions:`	This is a space-separated list of positions of cysteines in the heavy chain. This field is replaced by `CysPositions`: for scFv/scFab/XscFv
`HeavyDisulfidesIntra:`	This is a space-separated list of residue pairs involved in intra-chain disulfides in the heavy chain. The residue numbers in each residue pair are separated with a dash. In the case of fusion antibodies this record occurs for each instance (i.e. each antibody-like region) within the chain. In the case of scFv/scFab/XscFv (DART), this refers to the heavy chain part of the sequence. Where there are different disulfide isoforms, these

	are indicated with a [A], [B], etc in the same format as an instance indicator. (e.g. 11125)
`LightCysPositions:`	This is a space-separated list of positions of cysteines in the light chain. In the case of fusion antibodies this record occurs for each instance (i.e. each antibody-like region) within the chain. The record appears even if the antibody light chain like region is fused to a heavy chain. This field may be replaced by `CysPositions`: for scFv/scFab/XscFv
`CysPositions:`	Used with scFv/scFab/XscFv This is a space-separated list of positions of cysteines in the chain.
`LightDisulfidesIntra:`	This is a space-separated list of residue pairs involved in intra-chain disulfides in the light chain. The residue numbers in each residue pair are separated with a dash. In the case of fusion antibodies this record occurs for each instance (i.e. each antibody-like region) within the chain. The record appears even if the antibody light chain like region is fused to a heavy chain. In the case of scFv/scFab/XscFv (DART), this refers to the light chain part of the sequence. Where there are different disulfide isoforms, these are indicated with a [A], [B], etc in the same format as an instance indicator. (e.g. 11125)
`DisulfidesInterH1H2:`	This is a space-separated list of residue pairs involved in inter-chain disulfides between the two heavy chains. The residue numbers in each residue pair are separated with a dash. This record will be missing for scFvs, camelids, etc. Where there are different disulfide isoforms, these are indicated with a [A], [B], etc in the same format as an instance indicator. (e.g. 11125)
`DisulfidesInterL1H1``:`	This is a space-separated list of residue pairs involved in inter-chain disulfides between the first light chain and the first heavy chain. The residue numbers in each residue pair are separated with a dash. This record will be missing for truncated light chains and scFvs, and may be missing for Fab fragments. Where there are different disulfide isoforms, these are indicated with a [A], [B], etc in the same format as an instance indicator. (e.g. 11125)
`DisulfidesInterL2H2:`	This is a space-separated list of residue pairs

	involved in inter-chain disulfides between the second light chain and the second heavy chain. The residue numbers in each residue pair are separated with a dash. This record will be missing for truncated light chains, scFvs and Fab fragments. Where there are different disulfide isoforms, these are indicated with a [A], [B], etc in the same format as an instance indicator. (e.g. 11125)
`DisulfidesInter[x,y]:`	This is a more general case than `DisulfidesInterL1H1:` or `DisulfidesInterL2H2:` It consists of a space-separated list of residue pairs involved in inter-chain disulfides between chains not easily categorized as light or heavy (typically XscFVs (DARTs) or other interactions with scFvs or fusions), or in cases where there are multiple instances of light chains interacting with a VH fusion heavy chain or the same instance light chain interacting with multiple instances of heavy chains. Note that the specified residue pair (`a-b`) is in the order specified by `x,y` – i.e. a_x is bonded to b_y
`DisulfidesIntra[x]:`	This is a more general case than `LightDisulfidesIntra:` or `HeavyDisulfidesIntra:` It consists of a space-separated list of residue pairs involved in intra-chain disulfides within chains that are not light or heavy (typically a chain bonded to a fusion).
`HVGermline:`	This is the IMGT germline assignment for the heavy chain variable fragment. This is determined from the amino acids sequence and where different alleles result in the same sequence identity the allele with the lowest allele number is chosen. Where distal gene segments give the same score as a normal gene segment, the normal one is given. This record contains three fields: (1) the genus of the germline (e.g. *Homo*), (2) the species of the germline (e.g. *sapiens*), (3) the germline specifier (e.g. *IGHV3*01*)
`HJGermline:`	This is the IMGT germline assignment for the heavy chain joining fragment. This is determined from the amino acids sequence and where different alleles result in the same sequence identity the allele with the lowest allele number is chosen. Where distal gene segments give the same score as a normal gene segment, the normal one is given. This record contains three fields: (1) the genus of the germline (e.g. *Homo*), (2) the species of the germline (e.g. *sapiens*), (3) the germline specifier (e.g. *IGHJ1*01*)

`HCGermline:`	This is the IMGT germline assignment for the heavy chain constant fragment. This is determined from the amino acids sequence and where different alleles result in the same sequence identity the allele with the lowest allele number is chosen. Where distal gene segments give the same score as a normal gene segment, the normal one is given. This record contains three fields: (1) the genus of the germline (e.g. *Homo*), (2) the species of the germline (e.g. *sapiens*), (3) the germline specifier (e.g. *IGHG1*01*)
	Where the constant domains come from different germlines, each domain is indicated in parentheses. Where an instance qualifier is also used, this appears after the domain indicator. For example:
	`HCGermline(CH1): Homo sapiens IGHG2*01;`
	`HCGermline(CH2,CH3): Homo sapiens IGHG4*01;`
	`HCGermline(CH1)[1]: Homo sapiens IGHG1*01;`

`VHRange:`	This is a range of residues encoded by the V, D and J gene segments and equivalent to the protein VH domain. Sequential numbering is used and the residue numbers are separated by a hyphen (no spaces).
`CH1Range:`	This is a range of residues encoded by the C gene segment which is equivalent to the CH1 protein domain. Sequential numbering as observed in the final protein is used and the residue numbers are separated by a hyphen (no spaces). If there are any differences from the germline encoded sequence, these are indicated in parentheses using the form '(X125Y X130Y)' where X and Y are the 1-letter code representations of the germline encoded and observed amino acids respectively. If a germline encoded amino acid is deleted, this is indicated with '(X125-126del)' indicating that the expected 'X' residue between positions 125 and 126 is missing. If there is an insertion with respect to the germline, this is indicated with '(ins125Y)' Absent from scFvs.
`HingeRange:`	This is a range of residues encoded by the C gene segment which is equivalent to the hinge region in the protein. Sequential numbering as observed in the final protein is used and the residue numbers are separated by a hyphen (no spaces). If there are any differences from the germline

	encoded sequence, these are indicated in parentheses as described for CH1Range. Absent from scFvs and Fabs.
CH2Range:	This is a range of residues encoded by the C gene segment which is equivalent to the CH2 protein domain. Sequential numbering as observed in the final protein is used and the residue numbers are separated by a hyphen (no spaces). If there are any differences from the germline encoded sequence, these are indicated in parentheses as described for CH1Range. Absent from Fabs.
CH3Range:	This is a range of residues encoded by the C gene segment which is equivalent to the CH3 protein domain. Sequential numbering as observed in the final protein is used and the residue numbers are separated by a hyphen (no spaces). If there are any differences from the germline encoded sequence, these are indicated in parentheses as described for CH1Range. Absent from Fabs.
CH4Range:	This is a range of residues encoded by the C gene segment which is equivalent to the CH4 protein domain. Sequential numbering as observed in the final protein is used and the residue numbers are separated by a hyphen (no spaces). If there are any differences from the germline encoded sequence, these are indicated in parentheses as described for CH1Range. This record will only be present for IgM and IgE antibodies.

CHSRange:	This is a range of residues encoded by the C gene segment which is equivalent to the CHS region or the protein. This is normally the last two residues and the standard sequence is GK. If this is seen then the record will be shown as:
	`CHSRange: xxx-yyy;`
	Commonly the last lysine is deleted in which case the record is shown as:
	`CHSRange: xxx-xxx (Kyyydel);`
	Unusual final residues are indicated with
	`CHSRange: xxx-yyy (GxxxX KyyyX);`
	This record will be absent from all truncated chains unless just the CHS is absent in which case it will appear as:
	`CHSRange: (Gxxxdel Kyyydel);`

`MutationL:`	Records the purpose of mutations in the light chain. MutationL: NxxxM NxxxM NxxxM (purpose);
`MutationH:`	Records the purpose of mutations in the heavy chain. MutationH: NxxxM NxxxM NxxxM (purpose);
`Mutation[i]:`	Records the purpose of mutations in a chain specified by instance number. Mutation[1]: NxxxM NxxxM NxxxM (purpose);
`LVGermline:`	This is the IMGT germline assignment for the light chain variable fragment. This is determined from the amino acids sequence and where different alleles result in the same sequence identity the allele with the lowest allele number is chosen. Where distal gene segments give the same score as a normal gene segment, the normal one is given. This record contains three fields: (1) the genus of the germline (e.g. *Homo*), (2) the species of the germline (e.g. *sapiens*), (3) the germline specifier (e.g. *IGKV1*01*)
`LJGermline:`	This is the IMGT germline assignment for the light chain joining fragment. This is determined from the amino acids sequence and where different alleles result in the same sequence identity the allele with the lowest allele number is chosen. Where distal gene segments give the same score as a normal gene segment, the normal one is given. This record contains three fields: (1) the genus of the germline (e.g. *Homo*), (2) the species of the germline (e.g. *sapiens*), (3) the germline specifier (e.g. *IGKJ1*01*)
`LCGermline:`	This is the IMGT germline assignment for the light chain constant fragment. This is determined from the amino acids sequence and where different alleles result in the same sequence identity the allele with the lowest allele number is chosen. Where distal gene segments give the same score as a normal gene segment, the normal one is given. This record contains three fields: (1) the genus of the germline (e.g. *Homo*), (2) the species of the germline (e.g. *sapiens*), (3) the germline specifier (e.g. *IGKC*01*)
`VLRange:`	This is a range of residues encoded by the V and J gene segments which is equivalent to the VL protein domain. Sequential numbering as observed in the final protein is used and the residue numbers are separated by a hyphen (no spaces).
`CLRange``:`	This is a range of residues encoded by the C gene segment which is equivalent to the CL protein

	domain. Sequential numbering as observed in the final protein is used and the residue numbers are separated by a hyphen (no spaces). If there are any differences from the germline encoded sequence, these are indicated in parentheses as described for CH1Range. Absent from scFvs.
`CDRKabatL1:`	Specifies the sequence of CDR-L1 followed by the residue range in parentheses.
`CDRKabatL2:`	Specifies the sequence of CDR-L2 followed by the residue range in parentheses.
`CDRKabatL3:`	Specifies the sequence of CDR-L3 followed by the residue range in parentheses.
`CDRKabatH1:`	Specifies the sequence of CDR-H1 followed by the residue range in parentheses.
`CDRKabatH2:`	Specifies the sequence of CDR-H2 followed by the residue range in parentheses.
`CDRKabatH3:`	Specifies the sequence of CDR-H3 followed by the residue range in parentheses.

`CDRSource``:`	Allows the source of the CDRs to be specified for humanized antibodies. e.g.
	`CDRSource: human;`
	If different CDRs have come from different sources, then these may be specified in parentheses. e.g.
	`CDRSource(L1,L2,L3): human;`
	`CDRSource(H1,H2,H3): llama;`
	Instance numbers follow the parentheses. e.g.
	`CDRSource(H1,H2,H3)[1]: mouse;`

These records are followed by the sequence information which appears as

```
Heavy Chain:
//
```
and

```
Light Chain:
//
```
i.e. The sequence is introduced by an appropriate Chain record and is terminated by a pair of forward slashes.

Where there are multiple heavy or light chains, the instance number is included:

```
Heavy Chain[1]:
//
```

Where a chain occurs more than once in the complete compound (e.g. a light chain is shared between two heavy chains) the instance numbers are included as a comma-separated list:

```
Light Chain[1,2]:
//
```

Where a chain contains more than one instance (i.e. a fusion of some type, e.g. 10636), the instance numbers may be included as a hyphen-separated list for clarity:

```
Heavy Chain[4-5]:
//
```

In the case of scFv/scFab/XscFv (DART), where one chain is both heavy and light and cannot clearly be associated with one type or the other, the introductory record is simply:

```
Chain:
//
```

and where there are two chains (e.g. XscFv (DART)), then the instance qualifier is used:

```
Chain[1]:
//
Chain[2]:
//
```

As stated above, with the exception of the `Format:` record which must appear first, there is no requirement that records appear in a particular order, but the standard order as shown below should be used for standard antibodies. **Note that the lines may be concatenated into a single paragraph. There is no requirement for line breaks between the records.**

```
Format: ..... monoclonal antibody;
Type: ;
Conjugate: ;
Antigen: ;
HeavyChainClass: ;
HeavyChainLength: ;
LightChainClass: ;
LightChainLength: ;
HeavyPotentialNGlycos: ;
HeavyConfirmedNGlycos: ;
LightPotentialNGlycos: ;
LightConfirmedNGlycos: ;
HeavyCysPositions: ;
HeavyDisulfidesIntra: ;
LightCysPositions: ;
LightDisulfidesIntra: ;
```

```
DisulfidesInterH1H2: ;
DisulfidesInterL1H1: ;
DisulfidesInterL2H2: ;
HVGermline: ;
HJGermline: ;
HCGermline: ;
VHRange: ;
CH1Range: ;
HingeRange: ;
CH2Range: ;
CH3Range: ;
CHSRange: ;
LVGermline: ;
LJGermline: ;
LCGermline: ;
VLRange: ;
CLRange: ;
CDRKabatL1: ;
CDRKabatL2: ;
CDRKabatL3: ;
CDRKabatH1: ;
CDRKabatH2: ;
CDRKabatH3: ;
CDRSource: ;

Heavy Chain:
//

Light Chain:
//
```

Format for a fusion with another protein (e.g. 10500)

Fusions with another protein are indicated with a `Format:` record including the words “`fusion with`” followed by the species and name of the fusion protein(s).

The following records (used as required) indicate the range of the fusion:

FusionProteinHeavy:	Indicates a range of residues within a heavy chain that are part of a fusion protein. Instance qualifiers can be included as required. e.g.
	`FusionProteinHeavy: 443-970;`
	The range can also be followed by a multimerization indicator
	`(dimer|trimer|tetramer)`
	to indicate that the fusion protein forms a homo-multimer. e.g.
	`FusionProteinHeavy: 443-970 trimer;`

FusionProteinLight:	Indicates a range of residues within a light chain that are part of a fusion protein. Instance qualifiers can be included as required. The range can also be followed by a multimerization indicator to indicate that the fusion protein forms a homo-multimer.
Fusion:	Indicates instances that are fused together. e.g. `Fusion: 1 2;`
Interaction:	Indicates non-antibody domains that interact. i.e. when we have a fusion protein that consists of more than one chain. e.g. `Interaction: 1 3; (e.g.11029)`
FusionProtein:	Indicates a range of residues within a chain that can't be identified as light or heavy (i.e. where the sequence is indicated with a `Chain:` record) that are part of a fusion protein. Instance qualifiers can be included as required. The range can also be followed by a multimerization indicator to indicate that the fusion protein forms a homo-multimer. This may just be used as a more general case of FusionProteinHeavy: if the instance number is indicated and shown with Fusion: to be fused to an antibody heavy chain and the sequence then includes the instance numbers. (e.g. 11029)
FusionProteinHeavyDisulfides:	A list of ranges indicating disulfides in the fusion protein.
FusionProteinLightDisulfides:	A list of ranges indicating disulfides in the fusion protein.
FusionProteinDisulfides:	A list of ranges indicating disulfides in the fusion protein.
FusionProteinHeavyLinker:	A range indicating the linker between the heavy chain and a fusion protein
FusionProteinLightLinker:	A range indicating the linker between the light

	chain and a fusion protein

```
Format: xxxx monoclonal antibody fusion with human xxxx;
FusionProteinHeavy: ;
FusionProteinHeavyDisulfide: ;
Type: IgG1-kappa;
Antigen: ;
HeavyChainClass: ;
HeavyChainLength: ;
LightChainClass: ;
LightChainLength: ;
HeavyPotentialNGlycos: ;
HeavyConfirmedNGlycos:  ;
LightPotentialNGlycos: ;
LightConfirmedNGlycos: ;
HeavyCysPositions: ;
HeavyDisulfidesIntra: ;
LightCysPositions: ;
LightDisulfidesIntra: ;
DisulfidesInterH1H2: ;
DisulfidesInterL1H1: ;
DisulfidesInterL2H2: ;
HVGermline: ;
HJGermline: ;
HCGermline: ;
VHRange: ;
CH1Range: ;
HingeRange: ;
CH2Range: ;
CH3Range: ;
CHSRange: ;
LVGermline: ;
LJGermline: ;
LCGermline: ;
VLRange: ;
CLRange: ;
CDRKabatH1: ;
CDRKabatH2: ;
CDRKabatH3: ;
CDRKabatL1: ;
CDRKabatL2: ;
CDRKabatL3: ;

Heavy Chain:
//

Light Chain:
```

```
//
```

Format for an scFv

From PL119, the Format: record specifies that this is an scFv:

Format: humanized monoclonal scFv antibody;

The scFv contains two instances (1 and 2, representing the VL and VH) and a `Fusion:` record is used to indicate these are fused:

```
Fusion: 1 2;
```

The Domains[1,2]: record is used to clarify the order of the domains.

```
Domains[1,2]: VL VH;
```

An '`scFv`' qualifier to the `Type:` record is used to indicate an scFv:

```
Type[1]: Fv-kappa scFv;
Type[2]: Fv-heavy scFv;
```
(Fv-heavy was introduced from PL119)

A '`dimer`' qualifier to the `Type:` record is used to indicate scFv dimers:

```
Type[1]: Fv-kappa scFv dimer;
Type[2]: Fv-heavy scFv dimer;
```
(Fv-heavy was introduced from PL119)

A `Linker[1,2]:` record is used to indicate the linker:

```
Linker[1,2]: xxxxxxx;
```

The sequence will be indicated by a `Chain[1-2]` record. For example:

```
Format: humanized monoclonal scFv antibody;
Fusion: 1 2;
Antigen[1,2]: ;
Domains[1,2]: VL VH;
ChainLength[1,2]: ;

Type[1]: Fv-kappa scFv dimer;
LightChainClass[1]: ;
LightPotentialNGlycos[1]: ;
LightConfirmedNGlycos[1]: ;
LightCysPositions[1]: ;
LightDisulfidesIntra[1]: ;
LVGermline[1]: ;
LJGermline[1]: ;
VLRange[1]: ;
CDRKabatL1[1]: ;
```

```
CDRKabatL2[1]: ;
CDRKabatL3[1]: ;

Linker[1,2]: ;

Type[2]: Fv-heavy scFv dimer;
HeavyPotentialNGlycos[2]: ;
HeavyConfirmedNGlycos[2]: ;
HeavyCysPositions[2]: ;
HeavyDisulfidesIntra[2]: ;
HVGermline[2]: ;
HJGermline[2]: ;
VHRange[2]: ;
CDRKabatH1[2]: ;
CDRKabatH2[2]: ;
CDRKabatH3[2]: ;

Chain[1-2]:
//
```

Format for an scFv fusion with another protein (e.g. 10652, 10653)

The format is essentially the same as an scFv, but the fusion is indicated as a third instance. Relevant fields are:

```
Format: humanized monoclonal scFv antibody fusion with xxxx;
Fusion: 1 2 3;
Antigen[1,2]: xxxxx;
FusionProtein[3]: xxx-yyy;
Domains[1,2,3]: VL VH OTHER;
Chain[1-3-4]:
//
```

If the fusion protein forms a homo-multimer, this is indicated in the `FusionProtein[]:` record.

For example:

```
Format: humanized monoclonal antibody fusion with xxxx;
Fusion: 1 2 3;
Antigen[1,2]: ;
FusionProtein[3]: 254-386 trimer;
Domains[1,2,3]: VL VH OTHER;
ChainLength[1,2,3]: ;

Type[1]: Fv-kappa scFv dimer;
LightChainClass[1]: ;
LightPotentialNGlycos[1]: ;
LightConfirmedNGlycos[1]: ;
LightCysPositions[1]: ;
```

```
LightDisulfidesIntra[1]: ;
LVGermline[1]: ;
LJGermline[1]: ;
VLRange[1]: ;
CDRKabatL1[1]: ;
CDRKabatL2[1]: ;
CDRKabatL3[1]: ;

Linker[1,2]: ;

Type[2]: Fv-heavy scFv dimer;
HeavyPotentialNGlycos[2]: ;
HeavyConfirmedNGlycos[2]: ;
HeavyCysPositions[2]: ;
HeavyDisulfidesIntra[2]: ;
HVGermline[2]: ;
HJGermline[2]: ;
VHRange[2]: ;
CDRKabatH1[2]: ;
CDRKabatH2[2]: ;
CDRKabatH3[2]: ;

Chain[1-2-3]:
//
```

Format for a bispecific antibody (e.g. 10766)

In the case of a normal bispecific antibody there are two heavy chains and two light chains. The instance qualifier is used to indicate the light/heavy pair as appropriate. There will be a single Format: record indicating the overall antibody is a bispecific. e.g.:

```
Format: bispecific humanized monoclonal antibody;
```

Almost all remaining records will all be duplicated for the two instances; the exception being the `DisulfideInterH1H2` record which specifies the disulfide bonds between the first and second heavy chains. As stated above, with the exception of the `Format:` record which must appear first, there is no requirement that records appear in a particular order, but the standard order as shown below should be used for standard bispecific antibodies. The `DisulfideInterH1H2` record appears between the records for the first and second instance.

```
Format: bispecific .... monoclonal antibody;
Type[1]: ;
Conjugate[1]: ;
Antigen[1]: ;
HeavyChainClass[1]: ;
HeavyChainLength[1]: ;
LightChainClass[1]: ;
LightChainLength[1]: ;
```

```
HeavyPotentialNGlycos[1]: ;
HeavyConfirmedNGlycos[1]: ;
LightPotentialNGlycos[1]: ;
LightConfirmedNGlycos[1]: ;
HeavyCysPositions[1]: ;
HeavyDisulfidesIntra[1]: ;
LightCysPositions[1]: ;
LightDisulfidesIntra[1]: ;
DisulfidesInterL1H1: ;
HVGermline[1]: ;
HJGermline[1]: ;
HCGermline[1]: ;
VHRange[1]: ;
CH1Range[1]: ;
HingeRange[1]: ;
CH2Range[1]: ;
CH3Range[1]: ;
CHSRange[1]: ;
LVGermline[1]: ;
LJGermline[1]: ;
LCGermline[1]: ;
VLRange[1]: ;
CLRange[1]: ;
CDRKabatL1[1]: ;
CDRKabatL2[1]: ;
CDRKabatL3[1]: ;
CDRKabatH1[1]: ;
CDRKabatH2[1]: ;
CDRKabatH3[1]: ;
CDRSource[1]: ;

DisulfidesInterH1H2: ;

Type[2]: ;
Conjugate[2]: ;
Antigen[2]: ;
HeavyChainClass[2]: ;
HeavyChainLength[2]: ;
LightChainClass[2]: ;
LightChainLength[2]: ;
HeavyPotentialNGlycos[2]: ;
HeavyConfirmedNGlycos[2]: ;
LightPotentialNGlycos[2]: ;
LightConfirmedNGlycos[2]: ;
HeavyCysPositions[2]: ;
HeavyDisulfidesIntra[2]: ;
LightCysPositions[2]: ;
LightDisulfidesIntra[2]: ;
```

```
DisulfidesInterL2H2: ;
HVGermline[2]: ;
HJGermline[2]: ;
HCGermline[2]: ;
VHRange[2]: ;
CH1Range[2]: ;
HingeRange[2]: ;
CH2Range[2]: ;
CH3Range[2]: ;
CHSRange[2]: ;
LVGermline[2]: ;
LJGermline[2]: ;
LCGermline[2]: ;
VLRange[2]: ;
CLRange[2]: ;
CDRKabatL1[2]: ;
CDRKabatL2[2]: ;
CDRKabatL3[2]: ;
CDRKabatH1[2]: ;
CDRKabatH2[2]: ;
CDRKabatH3[2]: ;
CDRSource[2]: ;

Heavy Chain[1]:
//

Light Chain[1]:
//

Heavy Chain[2]:
//

Light Chain[2]:
//
```

Format for Bispecific Antibody scFv Fusions (e.g. 10431)

In the case of a bispecific antibody scFv fusion, there is a normal light chain, but the heavy chain has an scFv fused to the C-terminus. In other words the so-called heavy chain actually consists of a standard heavy chain with a linker joining it to a VL domain which in turn has a linker joining it to a VH domain. (The VH and VL domains of the scFv may also appear in the other order). There will be a single format record to indicate the overall nature of the protein:

```
Format: fusion ... monoclonal antibody;
```

Two additional record types are used for such constructs:

`Fusion:`	This is followed by a space separated list of instance numbers that are linked together in a fusion.
`Linker[x,y]:`	This indicates the range of residues that forms a linker between instances x and y.

An additional record may be needed for each part of the fusion if, for example, one part is human and the other humanized. This appears after the `Antigen[]:` records and the instance qualifiers will be the same as found in the `Antigen[]:` records.

`Source[x]:`	An optional field used with antibody/antibody fusions where the two parts are of different types (e.g. one is human and the other is humanized).

In this case the normal light chain has no instance qualifier since it part of a normal antibody and there will be two copies of the chain; thus the light chain records for the normal light chain will appear once with no instance qualifier. The absence of an instance qualifier can also be regarded as instance 0 (zero).

The so-called heavy chain contains three antibody-related regions and these are given instance numbers of 1, 2 and 3 – one of these is, of course, actually a light chain. This so-called heavy chain will be split into the three instances, two of which will be heavy chain records and one will be light chain records. A fusion record is used to indicate that these three instances are part of a single protein chain:

```
Fusion: 1 2 3;
```

An optional `Domains[x]:` record can be added to clarify the order of the domains in each chain.

```
Domains[0]: VL CL;
Domains[1,2,3]: VH CH1 CH2 CH3 VL VH;
```

Note that the `HeavyChainLength:` record refers to the length of the whole so-called heavy chain.

Linker records are used to indicate the residue ranges that form artificial linkers between the instances in the fusion chain. These need to specify the two instances that are fused. i.e.:

```
Linker[1,2]: www-xxx;
Linker[2,3]: yyy-zzz;
```

In the case of fused antibodies, the two antigens need to be tied to the respective instances which here refer to parts of chains rather than whole antibodies. Consequently the instance qualifier specifies both appropriate instances: `[0,1]` is used for the light and heavy pair that form the normal antibody and `[2,3]` for the scFv:

```
Antigen[0,1]: ;
```

If the fused antibodies are of different types, then this is indicated with the `Source:` record. This clarifies the information in the `Format:` record and appears after the `Antigen[]:` records and the instance qualifiers will be the same as found in the `Antigen[]:` records. For example:

```
Source[0,1]: human;
Source[2,3]: humanized;
```

Note that the light-chain region of the scFv fusion is annotated with `LightXXXX` records even though it is fused with the heavy chain. For the scFv region, the heavy-chain part of the fused sequence will have the records: `HeavyCysPositions, HeavyDisulfidesIntra, HVGermline, HJGermline,` VHRange while the light-chain part of the fused sequence will have the records: `LightChainClass, LightCysPositions, LightDisulfidesIntra, LVGermline, LJGermline, VLRange`.

As stated above, with the exception of the `Format:` record which must appear first, there is no requirement that records appear in a particular order, but the standard order as shown below should be used for standard bispecific antibodies. Note that the `DisulfideInterH1H2` record appears between the records for the first and second instance and the `Linker` records appear between the relevant instances.

```
Format: fusion ... monoclonal antibody;
Fusion: 1 2 3;
Type: ;
Antigen[0,1]: ;
Antigen[2,3]: ;

HeavyChainLength: ;

LightChainClass: ;
LightChainLength: ;
LightPotentialNGlycos: ;
LightConfirmedNGlycos: ;
LightCysPositions: ;
LightDisulfidesIntra: ;
LVGermline: ;
LJGermline: ;
10431LCGermline: ;
VLRange: ;
CLRange: ;
CDRKabatL1: ;
CDRKabatL2: ;
CDRKabatL3: ;
CDRSource(L1,L2,L3): ;
```

```
HeavyChainClass[1]: ;
HeavyPotentialNGlycos[1]: ;
HeavyConfirmedNGlycos[1]: ;
HeavyCysPositions[1]: ;
HeavyDisulfidesIntra[1]: ;
DisulfidesInterH1H2[1]: ;
DisulfidesInterL1H1[1]: ;
DisulfidesInterL2H2[1]: ;
HVGermline[1]: ;
HJGermline[1]: ;
HCGermline[1]: ;
VHRange[1]: ;
CH1Range[1]: ;
HingeRange[1]: ;
CH2Range[1]: ;
CH3Range[1]: ;
CHSRange[1]: ;
CDRKabatH1[1]: ;
CDRKabatH2[1]: ;
CDRKabatH3[1]: ;
CDRSource(H1,H2,H3)[1]: ;

Linker[1,2]: ;

HeavyCysPositions[2]: ;
HeavyDisulfidesIntra[2]: ;
HVGermline[2]: ;
HJGermline[2]: ;
VHRange[2]: ;
CDRKabatH1[2]: ;
CDRKabatH2[2]: ;
CDRKabatH3[2]: ;
CDRSource(H1,H2,H3)[2]: ;

Linker[2,3]: ;

LightChainClass[3]: ;
LightCysPositions[3]: ;
LightDisulfidesIntra[3]: ;
LVGermline[3]: ;
LJGermline[3]: ;
VLRange[3]: ;
CDRKabatL1[3]: ;
CDRKabatL2[3]: ;
CDRKabatL3[3]: ;
CDRSource(L1,L2,L3)[3]: ;
```

```
Heavy Chain[1-2-3]:
//

Light Chain[0]:
//
```

Format for an N-terminal VH/VL fusion with a normal antibody (e.g. 10622)

This is similar to the Bispecific Antibody scFv Fusion, but a VL is fused to the N-terminus of the light chain and a VH to the N-terminus of the heavy chain.

In this case 4 instances are defined: fusion-VL, light-chain, fusion-VH, heavy-chain.

Two `Fusion:` records are used to indicate the two fusions. An additional record, `Domains[x]:`, is used to list the domains in each instance:

`Domains[x]:`	Used with proteins such as CrossMabs to indicate (in order), the domains present in an instance or (together with the `Fusion:` records) a non-standard fusion protein chain. Allowed domains are: `VL, VH, CL, CH1, CH2, CH3, CH4, OTHER` (where `OTHER` represents non-antibody domains). Linkers are not listed.

```
Format: bispecific fusion xxx monoclonal antibody;

LightChainLength: ;
HeavyChainLength: ;

Fusion: 1 2;
Fusion: 3 4;
Domains[1]: VL;
Domains[2]: VL CL;
Domains[3]: VH;
Domains[4]: VH CH1 CH2 CH3;

Antigen[1,3]: ;
Antigen[2,4]: ;

LightCysPositions: ;
HeavyCysPositions: ;
DisulfidesInterL1H1[2,4]: ;
DisulfidesInterL2H2[2,4]: ;
DisulfidesInterH1H2[4]: ;

Type[1]: Fv-kappa;
LightChainClass[1]: kappa;
LightPotentialNGlycos[1]: ;
```

```
LightConfirmedNGlycos[1]: ;
LightDisulfidesIntra[1]: ;
LVGermline[1]: ;
LJGermline[1]: ;
VLRange[1]: ;
CDRKabatL1[1]: ;
CDRKabatL2[1]: ;
CDRKabatL3[1]: ;

Linker[1,2]: ;

Type[2]: IgGn-kappa;
LightChainClass[2]: kappa;
LightPotentialNGlycos[2]: ;
LightConfirmedNGlycos[2]: ;
LightDisulfidesIntra[2]: ;
LVGermline[2]: ;
LJGermline[2]: ;
LCGermline[2]: ;
VLRange[2]: ;
CLRange[2]: ;
CDRKabatL1[2]: ;
CDRKabatL2[2]: ;
CDRKabatL3[2]: ;

Type[3]: Fv-kappa;
HeavyPotentialNGlycos[3]: ;
HeavyConfirmedNGlycos[3]: ;
HeavyDisulfidesIntra[3]: ;
HVGermline[3]: ;
HJGermline[3]: ;
VHRange[3]: ;
CDRKabatH1[3]: ;
CDRKabatH2[3]: ;
CDRKabatH3[3]: ;

Linker[3,4]: ;

Type[4]: IgGn-kappa;
HeavyChainClass[4]: Gamma;
HeavyPotentialNGlycos[4]: ;
HeavyConfirmedNGlycos[4]: ;
HeavyDisulfidesIntra[4]: ;
HVGermline[4]: ;
HJGermline[4]: ;
HCGermline[4]: ;
VHRange[4]: ;
CH1Range[4]: ;
```

```
HingeRange[4]: ;
CH2Range[4]: ;
CH3Range[4]: ;
CHSRange[4]: ;
CDRKabatH1[4]: ;
CDRKabatH2[4]: ;
CDRKabatH3[4]: ;

Heavy Chain[3-4]:
//

Light Chain[1-2]:
//
```

Format for an scFv/xScFv (DART) (e.g. 10569)

(Format: record changed from PL119; Fv-heavy introduced from PL119)

```
Format: humanized monoclonal scFv antibody;

Fusion: 1 2;
Fusion: 3 4;
Antigen[1,4]: ;
Antigen[2,3]: ;

ChainLength[1,2]: ;

Type[1]: Fv-lambda XscFv;
LightChainClass[1]: ;
LightPotentialNGlycos[1]: ;
LightConfirmedNGlycos[1]: ;
LightCysPositions[1]: ;
LightDisulfidesIntra[1]: ;
LVGermline[1]: ;
LJGermline[1]: ;
VLRange[1]: ;
CDRKabatL1[1]: ;
CDRKabatL2[1]: ;
CDRKabatL3[1]: ;

Linker[1,2]: ;

Type[2]: Fv-heavy XscFv;
HeavyChainClass[2]: ;
HeavyPotentialNGlycos[2]: ;
HeavyConfirmedNGlycos[2]: ;
HeavyCysPositions[2]: ;
HeavyDisulfidesIntra[2]: ;
HVGermline[2]: '
```

```
HJGermline[2]: ;
VHRange[2]: ;
CDRKabatH1[2]: ;
CDRKabatH2[2]: ;
CDRKabatH3[2]: ;

DisulfidesInter[2,4]: ;

ChainLength[3,4]: ;

Type[3]: Fv-kappa XscFv;
LightChainClass[3]: ;
LightPotentialNGlycos[3]: ;
LightConfirmedNGlycos[3]: ;
LightCysPositions[3]: ;
LightDisulfidesIntra[3]: ;
LVGermline[3]: ;
LJGermline[3]: ;
VLRange[3]: ;
CDRKabatL1[3]: ;
CDRKabatL2[3]: ;
CDRKabatL3[3]: ;

Linker[3,4]: ;

Type[4]: Fv-heavy XscFv;
HeavyChainClass[4]: ;
HeavyPotentialNGlycos[4]: ;
HeavyConfirmedNGlycos[4]: ;
HeavyCysPositions[4]: ;
HeavyDisulfidesIntra[4]: ;
HVGermline[4]: ;
HJGermline[4]: ;
VHRange[4]: ;
CDRKabatH1[4]: ;
CDRKabatH2[4]: ;
CDRKabatH3[4]: ;

Chain[1-2]:
//

Chain[3-4]:
//
```

Format for a Bispecific CrossMab (e.g. 10563)

In a bispecific Crossmab, one of the two antibodies swaps the CL/CK and CH1 domains, so the light chain contains VL-CH1 and the heavy chain contains VH-CL-CH2-CH3.

The crossed chain adds the keyword `Crossmab` to the end of the `Type[]:` record. `Domains[]:` records are also included to indicate the order of domains.

```
Format: bispecific xxxx monoclonal antibody;

Domains[1]: VH CH1 CH2 CH3;
Domains[2]: VL CL;
Domains[3]: VH CL CH2 CH3;
Domains[4]: VL CH1;

Antigen[1,2]: ;
DisulfidesInter[1,2]: ;

Type[1]: IgG1-kappa;

HeavyChainClass[1]: ;
HeavyChainLength[1]: ;
HeavyPotentialNGlycos[1]: ;
HeavyConfirmedNGlycos[1]: ;
HeavyCysPositions[1]: ;
HeavyDisulfidesIntra[1]: ;
HVGermline[1]: ;
HJGermline[1]: ;
HCGermline[1]: ;
VHRange[1]: ;
CH1Range[1]: ;
HingeRange[1]: ;
CH2Range[1]: ;
CH3Range[1]: ;
CHSRange[1]: ;
CDRKabatH1[1]: ;
CDRKabatH2[1]: ;
CDRKabatH3[1]: ;

Type[2]: IgG1-kappa;

LightChainClass[2]: ;
LightChainLength[2]: ;
LightPotentialNGlycos[2]: ;
LightConfirmedNGlycos[2]: ;
LightCysPositions[2]: ;
LightDisulfidesIntra[2]: ;
```

```
LVGermline[2]: ;
LJGermline[2]: ;
LCGermline[2]: ;
VLRange[2]: ;
CLRange[2]: ;
CDRKabatL1[2]: ;
CDRKabatL2[2]: ;
CDRKabatL3[2]: ;

DisulfidesInter[1,3]: ;

Antigen[3,4]: ;
DisulfidesInter[3,4]: ;

Type[3]: IgG1-lambda Crossmab;
HeavyChainClass[3]: ;
HeavyChainLength[3]: ;
HeavyPotentialNGlycos[3]: ;
HeavyConfirmedNGlycos[3]: ;
HeavyCysPositions[3]: ;
HeavyDisulfidesIntra[3]: ;
HVGermline[3]: ;
HJGermline[3]: ;
LCGermline(CL)[3]: ;
HCGermline(CH2,CH3)[3]: ;
VHRange[3]: ;
CLRange[3]: ;
HingeRange[3]: ;
CH2Range[3]: ;
CH3Range[3]: ;
CHSRange[3]: ;
CDRKabatH1[3]: ;
CDRKabatH2[3]: ;
CDRKabatH3[3]: ;

Type[4]: IgG1-lambda Crossmab;
LightChainClass[4]: ;
LightChainLength[4]: ;
LightPotentialNGlycos[4]: ;
LightConfirmedNGlycos[4]: ;
LightCysPositions[4]: ;
LightDisulfidesIntra[4]: ;
LVGermline[4]: ;
LJGermline[4]: ;
HCGermline(CH1)[4]: ;
VLRange[4]: ;
CH1Range[4]: ;
CDRKabatL1[4]: ;
```

```
CDRKabatL2[4]: ;
CDRKabatL3[4]: ;

Heavy Chain[1]:
//
Light Chain[2]:
//
Heavy Chain[3]:
//
Light Chain[4]:
//
```

Format for a Chimeric Antigen Receptor (scFv-TCRa / TCRb)

The CAR format includes a non-antibody chain (the TCR beta chain) that is paired with the fusion (TCR alfa chain) joined to the scFv.

```
Format: humanized monoclonal antibody fusion with human T-cell
receptor;

Fusion: 1 2 3;
Domains[1,2,3]: VL VH OTHER;
ChainLength[1,2,3]: ;

Antigen[1,2]: ;

Type[1]: Fv-kappa scFv;
LightChainClass[1]: ;
LightPotentialNGlycos[1]: ;
LightConfirmedNGlycos[1]: ;
LightCysPositions[1]: ;
LightDisulfidesIntra[1]: ;
LVGermline[1]: ;
LJGermline[1]: ;
VLRange[1]: ;
CDRKabatL1[1]: ;
CDRKabatL2[1]: ;
CDRKabatL3[1]: ;

Linker[1,2]: ;

Type[2]: Fv-kappa scFv;
HeavyPotentialNGlycos[2]: ;
HeavyConfirmedNGlycos[2]: ;
HeavyCysPositions[2]: ;
HeavyDisulfidesIntra[2]: ;
HVGermline[2]: ;
```

```
HJGermline[2]: ;
VHRange[2]: ;
CDRKabatH1[2]: ;
CDRKabatH2[2]: ;
CDRKabatH3[2]: ;

Linker[2,3]: ;

Type[3]: OTHER;
CysPositions[3]: ;
FusionProtein[3]: ;
FusionProteinDisulfides[3]: ;
DisulfidesInter[3,4]: ;
FusionProteinDisulfides[4]: ;

Type[4]: OTHER;
CysPositions[4]: ;
DisulfidesIntra[4]: ;
ChainLength[4]: ;
PotentialNGlycos[4]: ;
ConfirmedNGlycos[4]: ;

Chain[1-2-3]:
//

Chain[4]:
//
```

Format for an Fc fused to another protein

e.g. 10661, 10695

```
Request:

Format: human monoclonal antibody Fc fusion with human xxxx;
Fusion: 1 2;
Domains[1]: CH2 CH3;
Domains[2]: OTHER;

HeavyChainLength: ;

Type[1]: Fc;
HeavyChainClass[1]: Gamma1;
HeavyPotentialNGlycos[1]: ;
HeavyConfirmedNGlycos[1]: ;
HeavyCysPositions[1]: ;
HeavyDisulfidesIntra[1]: ;
```

```
DisulfidesInterH1H2[1]: ;
HCGermline[1]: ;
CH2Range[1]: ;
CH3Range[1]: ;
CHSRange[1]: ;

Linker[1,2]: ;

Type[2]: OTHER;
FusionProtein[2]: ;
CysPositions[2]: ;
DisulfidesIntra[2]: ;
PotentialNGlycos[3]: ;

Heavy Chain[1-2]:
//
```

Format for a VHFc (single domain antibody fused to an Fc, e.g. 10930)

```
Request: 10930

Format: humanized monoclonal VHFc antibody;
Type: VHFc;
Antigen: ;
HeavyChainClass: ;
HeavyChainLength: ;
HeavyPotentialNGlycos: ;
HeavyConfirmedNGlycos: ;
HeavyCysPositions: ;
HeavyDisulfidesIntra: ;
DisulfidesInterH1H2: ;
HVGermline: ;
HJGermline: ;
HCGermline: ;
VHRange: ;
HingeRange: ;
CH2Range: ;
CH3Range: ;
CHSRange: ;
CDRKabatH1: ;
CDRKabatH2: ;
CDRKabatH3: ;
CDRSource: camelid;

Heavy Chain:
//
```