determine the optimal molecule-specific chain expression ratio with high throughput. ModVec vector library provide a platform to rapidly determine the best expression configuration for individual multichain molecule. ModVec can be used in more generalized applications to assemble complex large DNA molecules for purposes other than expression vectors for antibody therapeutics.

SUMMARY

[0005] In one aspect, the present disclosure is directed to a method for expressing a multi-chain protein comprising at least two different polypeptide chains comprising

[0006] a. providing polynucleotide sequences that encode for the at least two different polypeptides, wherein said polynucleotide sequences are flanked on both 5' and 3' sides by a type IIS restriction endonuclease cleavage site followed by a recognition site thereof;

[0007] b. providing two or more sets of element sequences; wherein each set of element sequences together comprise at least one functional expression cassette,

[0008] wherein each element sequence is flanked on both 5' and 3' sides by a type IIS restriction endonuclease cleavage site followed by a recognition site thereof,

[0009] wherein the type IIS restriction endonuclease recognition sites and cleavage sites are selected so that the sets of element sequences may be assembled into a functional expression cassette;

[0010] c. providing at least one backbone entry vector, wherein the backbone entry vector comprises a plasmid comprising at least one type IIS restriction endonuclease cleavage site followed by a recognition site thereof and sequences for maintenance of the vector in bacterial cells;

[0011] d. assembling the polynucleotide sequences that encode for the at least two different polypeptides with the two or more sets of element sequences to generate functional expression cassettes in the backbone entry vector, using a method based on the use of type IIS restriction enzyme digestion and ligation via the cleavage sites and overhangs resulting in ligated vectors comprising at least two functional expression cassettes that are capable of expressing the at least two different polypeptide chains;

[0012] e. transfecting mammalian cells with the ligated vectors;

[0013] f. comparing the levels of preferred product produced and levels of impurities of each of the at least two polypeptides; and

[0014] g. identifying the optimal expression ratio of the two polypeptides for the transfected mammalian cells and identifying the transfected mammalian cells that have the optimal expression ratio of the at least two polypeptide chains.

[0015] In one example, the method further comprises:

[0016] h. identifying the arrangement of vector elements that provides the optimal expression ratio of the at least two polypeptide chains.

[0017] In one aspect, the present disclosure is directed to a vector comprising the arrangement of vector elements that provides the optimal expression ratio of the at least two polypeptide chains.

[0018] In one example, the element sequences comprise at least two promoter sequences and at least two polyA sequences.

[0019] In one example, the levels of preferred product produced are measured by a method selected from the group consisting of cation exchange chromatography (reduced or non-reduced), mass spectrometry, any other chromatographic separation, or a combination thereof.

[0020] In one example, none of the type IIS restriction endonuclease cleavage sites produce 5' four nucleotide overhangs selected from the group consisting of GTAA, TCCA, and CACA upon cleavage by the type IIS restriction endonuclease.

[0021] In one example, the type IIS restriction endonuclease cleavage site is selected from the group consisting of:

[0022] AGGT, AGTA, ATCA, CAGT, CCAT, GAAT, GAGG, GGCA, GGTC, TAGC, TCTT, GGAG, and CCAC.

[0023] In one example, the mammalian cells are selected from the group consisting of CHO cells, CHOK1 cells, DXB-11, DG-44, COS-7, HEK293, BHK, TM4, CV1, VERO-76, HELA, MDCK, BRL 3A, W138, Hep G2, MMT 060562, TRI cells, MRC 5 cells, FS4 cells, and mammalian myeloma cells.

[0024] In one example, the optimal expression ratio of the at least two polypeptides is selected from the group consisting of 1:1, 1:2, and 1:3.

[0025] In one example, the multi-chain protein comprises a first antibody heavy chain, a first antibody light chain, a second antibody heavy chain, and a second antibody light chain, wherein the first antibody heavy chain associates with the first antibody light chain to bind a first antigen or epitope and the second antibody heavy chain associates with the second antibody light chain to bind a second antigen or epitope, wherein the optimal expression ratio of the first antibody heavy chain, the first antibody light chain, the second antibody heavy chain, and the second antibody light chain is 1:1:1:1.

[0026] In one example, the multi-chain protein comprises a first antibody heavy chain, a second antibody heavy chain, and a common antibody light chain, wherein the first antibody heavy chain associates with the common antibody light chain to bind a first antigen or epitope and the second antibody heavy chain associates with the common antibody light chain to bind a second antigen or epitope, wherein the optimal expression ratio of the first antibody heavy chain, the second antibody heavy chain, and the common antibody light chain is 1:1:2.

[0027] In one example, the multi-chain protein comprises an antibody heavy chain, a first antibody light chain, a modified antibody heavy chain, and a second antibody light chain,

[0028] wherein the modified antibody heavy chain comprises, N-terminal to C-terminal, one of the following structures selected from the following group:

[0029] VH-CH1-binding domain (BD)-hinge-CH2-CH3;

[0030] BD-VH-CH1-hinge-CH2-CH3; and

[0031] VH-CH1-hinge-CH2-CH3-BD;

[0032] wherein the BD is selected either a single-chain Fv (scFv) or a single-chain Fab (scFab);

[0033] wherein the antibody heavy chain associates with the first antibody light chain to bind a first antigen or epitope and the VH of the modified antibody heavy chain associates with the second antibody light chain to bind a second antigen or epitope, wherein the BD binds to a third antigen or epitope, and

[0034] wherein the optimal expression ratio of the first antibody heavy chain, the first light chain, the modified antibody heavy chain, and the second antibody light chain is 1:1:1:1.

[0035] In one example, the multi-chain protein comprises an antibody heavy chain, a modified antibody heavy chain, and a common antibody light chain,

[0036] wherein the modified antibody heavy chain comprises, N-terminal to C-terminal, one of the following structures selected from the following group:

[0037] VH-CH1-binding domain (BD)-hinge-CH2-CH3;

[0038] BD-VH-CH1-hinge-CH2-CH3; and

[0039] VH-CH1-hinge-CH2-CH3-BD;

[0040] wherein the BD is selected either a single-chain Fv (scFv) or a single-chain Fab (scFab);

[0041] wherein the antibody heavy chain associates with the common antibody light chain to bind a first antigen or epitope and the VH of the modified antibody heavy chain associates with the common antibody light chain to bind the first antigen or epitope, wherein the BD binds to a second antigen or epitope, and

[0042] wherein the optimal expression ratio of the first antibody heavy chain, the modified antibody heavy chain, and the common antibody light chain is 1:1:2.

BRIEF DESCRIPTION OF THE DRAWINGS

[0043] Figures 1A-1E depicts the Modular vector platform for high throughput vector engineering ("ModVec"). Figure 1A) Example sequence of DNA module. Each module (depicted in Ns and grey shade) is flanked by Golden Gate adaptors (boxed region) which contain recognition site for *BsmBI* (highlighted in bold). *BsmBI* cut sites are marked with ^ and overhangs generated through *BsmI* digestion are underlined. Figure 1B) Assembly of an expression vector is achieved through Golden Gate reaction