(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
5 January 2012 (05.01.2012)

PCT

(10) International Publication Number
WO 2012/003277 A2

(51) **International Patent Classification:**
A61K31/047 (2006.01)

(21) **International Application Number:**
PCT/US201 1/042529

(22) **International Filing Date:**
30 June 201 1 (30.06.201 1)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**
61/360,040 30 June 2010 (30.06.2010) US

(71) **Applicant** *(for all designated States except US):* **CODEXIS, INC.** [US/US]; 200 Penobscot Drive, Redwood City, California 94063 (US).

(72) **Inventor; and**
(75) **Inventor/Applicant** *(for US only):* **ALVIZO, Oscar** [US/US]; 200 Penobscot Drive, Redwood City, California 94063 (US).

(74) **Agents: WHITING, Adam, K.** et al; Codexis, Inc., 200 Penobscot Drive, Redwood City, California 94063 (US).

(81) **Designated States** *(unless otherwise indicated, for every kind of national protection available):* AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PE, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) **Designated States** *(unless otherwise indicated, for every kind of regional protection available):* ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

— *with sequence listing part of description (Rule 5.2(a))*

WO 2012/003277 A2

(54) **Title:** HIGHLY STABLE BETA-CLASS CARBONIC ANHYDRASES USEFUL IN CARBON CAPTURE SYSTEMS

(57) **Abstract:** The present disclosure relates to β-class carbonic anhydrase polypeptides having improved properties including increased thermostability and/or stability in the presence of amine compounds, ammonia, or carbonate ion. The present disclosure also provides formulations and uses of the polypeptides for accelerating the absorption of carbon dioxide from a gas stream into a solution as well as for the release of the absorbed carbon dioxide for further treatment and/or sequestering. Also provided are polynucleotides encoding the carbonic anhydrase polypeptides and host cells capable of expressing them.

HIGHLY STABLE BETA-CLASS CARBONIC ANHYDRASES USEFUL IN CARBON CAPTURE SYSTEMS

1. TECHNICAL FIELD

[0001] The present disclosure relates to beta-class carbonic anhydrase polypeptides that exhibit high solvent and thermostability, and methods of using these polypeptides in carbon capture systems.

2. REFERENCE TO SEQUENCE LISTING, TABLE OR COMPUTER PROGRAM

[0002] The Sequence Listing concurrently submitted electronically under 37 C.F.R. § 1.821 *via* EFS-Web in a computer readable form (CRF) as file name CX4-061USPl_ST25.txt is herein incorporated by reference. The electronic copy of the Sequence Listing was created on June 30, 2010 with a file size of 430 kilobytes.

3. BACKGROUND

[0003] The enzyme, carbonic anhydrase ("CA") (EC 4.2.1.1), catalyzes the reversible reactions depicted in Scheme 1:

Scheme 1

$$C0_2 + H_20 \overset{CA}{\rightleftharpoons} HC0_3^- + H^+$$

[0004] In the forward or "hydration" reaction, CA combines carbon dioxide and water to provide bicarbonate and a proton, or depending on the pH, to provide carbonate (CO_3^{-2}) and two protons. In the reverse, or "dehydration" reaction, CA combines bicarbonate and a proton to provide carbon dioxide and water. Carbonic anhydrases are metalloenzymes that typically have Zn^{+2} in the active site. However carbonic anhydrases having *e.g.* Co^{+2} or Cd^{+2} in the active site have been reported. At least three classes of carbonic anhydrases have been identified in nature.

[0005] The a-class carbonic anhydrases are found in vertebrates, bacteria, algae, and the cytoplasm of green plants. Vertebrate a-carbonic anhydrases are among the fastest enzymes known, exhibiting a turnover number (k_{cat}) (the number of molecules of substrate converted by an enzyme to product per catalytic site per unit of time) of 10^6 sec^{-1}. The β-class carbonic anhydrases are found in bacteria, algae, and chloroplasts, while γ-class carbonic anhydrases are found in *Archaea* and some bacteria. Although carbonic anhydrases of each of these classes have similar active sites, they do not exhibit significant overall amino acid sequence homology and they are structurally distinguishable from one another. Hence, these three classes of carbonic anhydrase provide an example of convergent evolution.

[0006] It has been proposed that carbonic anhydrases could be used as a biological catalyst to accelerate the capture of carbon dioxide produced by produced by combustion of fossil fuels. *See e.g.*, U.S. Pat. Nos. 6,143,556, 6,524,843 B2, 7,176,017 B2, 7,596,952 B2, 7,579,185 B2, 7,740,689

B2, 7,132,090 B2; U.S. Pat. Publ. Nos. 2009/0155889A1, 2010/0086983A1; PCT Publ. Nos. WO2006/089423A1, WO2010/014773A1, WO2010/045689A1. State-of-the-art carbonic anhydrases, however, are not well-suited for use in such applications because of their relative lack of stability and/or activity under the process conditions required. Accordingly, there is a need in the art for carbonic anhydrases that can effectively accelerate the absorption of carbon dioxide from a gas stream and/or later accelerate desorption of the carbon dioxide from the capture solution under process relevant conditions, such as presence in solution with high concentrations of other $C0_2$ absorption mediating compounds *(e.g.,* amines, ammonia, carbonate ions), elevated temperatures *(e.g.,* 40°C or above, or 15°C or below in NH_3), alkaline pHs *(e.g.,* pH 8-12), and extended periods of exposure to these challenging conditions *(e.g.,* days to weeks). In addition, such carbonic anhydrases should also be stable to variations in these conditions, *e.g.* stable not only at a relatively alkaline pH suitable for hydration and sequestration of carbon dioxide but also at a relatively acidic pH suitable for subsequent release and/or recapture of the hydrated and/or sequestered carbon dioxide.

4. SUMMARY

[0007] The present disclosure provides carbonic anhydrase polypeptides that are capable of accelerating the absorption of carbon dioxide from a gas stream into a solution under suitable conditions useful for various carbon capture processes *(e.g.,* amine-based compounds, chilled ammonia) and have increased thermal and/or solvent stability relative to known beta-class carbonic anhydrases. In particular embodiments, the disclosure provides recombinant (or engineered) carbonic anhydrase polypeptides that are capable of hydrating carbon dioxide to bicarbonate with increased activity relative to the wild-type enzyme from *Desulfovibrio vulgaris* after exposure to high concentrations of amine solvent, ammonia solvent, or carbonate ion, and/or any of these solvents and substantially increased and/or decreased temperatures *(e.g.,* $T > 40°C$ and/or $T < 20°C$). Additionally, the present disclosure provides polynucleotides encoding the carbonic anhydrase polypeptides, methods and hosts cells for the expression of those polypeptides, as well as methods, processes, and bioreactors for using the presently disclosed polypeptides.

[0008] In one aspect, the carbonic anhydrase polypeptides described herein have an amino acid sequence that has one or more amino acid differences as compared to a wild-type carbonic anhydrase or an engineered carbonic anhydrase that result in an improved property of the enzyme. Generally, the engineered carbonic anhydrase polypeptides have an improved property as compared to the wild-type carbonic anhydrase enzyme obtained from *Desulfovibrio vulgaris* str. "Miyazaki F" (GenBank accession ACL09337.1 GI:2 18758438; SEQ ID NO: 2). Improvements in an enzyme property include increases in thermostability, solvent stability, increased level of expression, enzyme activity at elevated pH, and enzyme stability and/or activity during pH variations, as well as reduced inhibition by solvents and/or compounds in solution *(e.g.,* product inhibition by carbonate or bicarbonate). Improvements in an enzyme property of the recombinant carbonic anhydrases disclosed herein also

include increased stability in the presence in the enzyme solution of compounds that mediate the absorption or sequestration of carbon dioxide, including, for example, ammonia, carbonate ions, amine compounds *(e.g.,* monoethanolamine (MEA), methyldiethanolamine (MDEA), 2-aminomethylpropanolamine (AMP), 2-(2-aminoethylamino)ethanol (AEE), triethanolamine (TEA), 2-amino-2-hydroxymethyl-l,3-propanediol (AHPD), piperazine, piperidine, mono- and diethanolamine) .

[0009] In certain embodiments, the present disclosure provides a recombinant carbonic anhydrase polypeptide having an improved enzyme property relative to the reference sequence of SEQ ID NO:2, wherein the polypeptide comprises an amino acid sequence at least about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical SEQ ID NO:2, and an amino acid difference at one or more of the following positions relative to SEQ ID NO:2: X2; X4; X15; X16; X22; X30; X31; X32; X34; X35; X37; X40; X42; X43; X47; X56; X60; X68; X70; X84; X86; X95; X96; X97; X I 19; X120; X121; X124; X131; X138; X139; X142; X143; X144; X145; X147; X148; X157; X159; X168; X170; X178; X200; X207; X213; X219; X221; X222; and X223. In certain embodiments, the recombinant carbonic anhydrase polypeptide comprises an amino acid sequence having one or more of the following amino acid differences relative to SEQ ID NO: 2: X2R; X2T; X4F; X4M; X15R; X16S; X22G; X30A; X30K; X30L; X30Q; X30R; X31P; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40Q; X40W; X42A; X43M; X43V; X47R; X56S; X60C; X60V; X68A; X68G; X70I; X84K; X84N; X84Q; X84R; X84S; X86A; X95V; X96A; X96C; X96E; X96K; X97F; X I 19K; X I 19L; X I 19M; X I 19T; X120R; X121H; X121K; X121L; X121Q; X121T; X121V; X121W; X124G; X124R; X131L; X138F; X138L; X138W; X139H; X139K; X139M; X139Q; X142L; X143M; X144A; X144L; X145C; X145F; X145W; X147E; X147F; X147G; X147T; X148A; X148T; X157A; X159H; X159R; X159V; X168E; X170F; X178G; X200R; X207E; X207N; X213E; X213Q; X219T; X221C; X222C; X223C; andX223Q.

[0010] In certain embodiments, the disclosure provides exemplary recombinant carbonic anhydrase polypeptides having an improved enzyme property relative to a reference polypeptide of SEQ ID NO:2 which comprise an amino acid sequence selected from the group consisting of SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 240, 242, 244, 246, 248, 250, 252, 254, 256, 258, 260, 262, 264, 266, and 268.

[0011] In some embodiments, the disclosure provides a recombinant carbonic anhydrase which has an amino acid sequence that is at least about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%,

79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the amino acid sequence corresponding to SEQ ID NO: 2, and further comprises any one or more of the amino acid differences, or combinations of differences (see *e.g.,* amino acid differences presented in Tables 2A and 2B), found in any of the exemplary recombinant carbonic anhydrase polypeptide amino acid sequences. In some embodiments, these recombinant carbonic anhydrase polypeptides can have one or more amino acid differences at other residue positions, and/or insertions, deletions at other positions, and/or additional amino or carboxy terminal extensions.

[0012] In some embodiments, the recombinant carbonic anhydrase polypeptides of the present disclosure exhibit the improved property of increased stability as compared to the wild-type enzyme of SEQ ID NO: 2, wherein the increased stability is determined under specified suitable conditions, including but not limited to conditions of temperature, solution composition *(e.g.,* presence in solution of specified concentration an amine compound, ammonia, and/or carbonate ion) solution CO_2 loading *(e.g., a* = 0 to about 0.7), and/or exposure to stability challenge conditions *(e.g.,* increased temperature or decreased temperature).

[0013] In some embodiments, the recombinant carbonic anhydrase polypeptides exhibit at least 1.2-fold, at least 1.3-fold, at least 1.5-fold, at least 2-fold, at least 3-fold, at least 4-fold, at least 5-fold, at least 10-fold, or at least 25-fold, or more, increased stability as compared to a reference polypeptide *(e.g.,* wild-type of SEQ ID NO: 2, or a recombinant carbonic anhydrase polypeptide of SEQ ID NO: 26) under suitable conditions. In some embodiments, the increased stability is determined by measuring post-challenge relative activity of hydrating carbon dioxide or dehydrating bicarbonate.

[0014] In some embodiments, the carbonic anhydrase polypeptides of the disclosure are improved as compared to SEQ ID NO: 2 with respect to their stability in the presence of an amine compound, ammonia, or carbonate ion in the enzyme solution.

[0015] In some embodiments, the recombinant carbonic anhydrase polypeptides of the present disclosure exhibit the improved property of increased thermostability - *e.g.,* retain increased activity relative to the polypeptide of SEQ ID NO: 2 after exposure to challenge conditions comprising an elevated temperature within the range of from about 30°C to about 100°C, from about 40°C to about 80°C, or within a range of from 40°C to about 60°C.

[0016] In another aspect, the present disclosure provides polynucleotides encoding the engineered carbonic anhydrases described herein or polynucleotides that hybridize to such polynucleotides under highly stringent conditions. The polynucleotide can include promoters and other regulatory elements useful for expression of the encoded engineered carbonic anhydrases, and can utilize codons optimized for specific desired expression systems. In some embodiments, the polynucleotides encode a carbonic anhydrase polypeptide having at least the following amino acid sequence as compared to

the amino acid sequence of SEQ ID NO: 2, and further comprising one or more amino acid differences selected from the group of amino acid differences provided in Tables 2A and 2B. Exemplary polynucleotides include, but are not limited to, a polynucleotide sequence of any of SEQ ID NO: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 181, 183, 185, 187, 189, 191, 193, 195, 197, 199, 201, 203, 205, 207, 209, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 239, 241, 243, 245, 247, 249, 251, 253, 255, 257, 259, 261, 263, 265, 267, 269, 271, 273, 275, 277, 279, 281, 283, and 285.

[0017] In another aspect, the present disclosure provides host cells comprising the polynucleotides and/or expression vectors described herein. The host cells may be *D. vulgaris* or they may be a different organism, such as *E. coli, Saccharomyces cerevisiae, Bacillus spp. (e.g., B. amyloliquefaciens, B. licheniformis, B. megaterium, B. stearothermophilus,* and *B. subtilis), Streptomyces spp. (e.g., Streptomyces coelicolor* (DSM 40680) and *Streptomyces lividans* (DSM 40434)), or filamentous fungal organisms such as *Aspergillus spp.* including but not limited to *A. niger, A. nidulans, A. awamori, A. oryzae, A. sojae and A. kawachi; Trichoderma reesei; Chrysosporium lucknowense; Myceliophthora thermophilia; Fusarium venenatum; Neurospora crassa; Humicola insolens; Humicola grisea; Penicillum verruculosum; Thielavia terrestris;* and teleomorphs, or anamorphs and synonyms or taxonomic equivalents thereof. The host cells can be used for the expression and isolation of the recombinant carbonic anhydrase enzymes described herein, or, alternatively, they can be used directly for carrying out the reactions of Scheme 1.

[0018] In some embodiments, the disclosure provides a method of producing a recombinant carbonic anhydrase polypeptide of the present disclosure, wherein said method comprises the steps of: (a) transforming a host cell with an expression vector polynucleotide encoding the recombinant carbonic anhydrase polypeptide; (b) culturing said transformed host cell under conditions whereby said recombinant carbonic anhydrase polypeptide is produced by said host cell; and (c) recovering said recombinant carbonic anhydrase polypeptide from said host cells. In some embodiments, the method of producing the recombinant carbonic anhydrase may be carried out wherein said expression vector comprises a secretion signal, and said cell is cultured under conditions whereby the recombinant carbonic anhydrase polypeptide is secreted from the cell. In some embodiments of the method, the expression vector comprises a polynucleotide encoding a secretion signal.

[0019] In some embodiments, the carbonic anhydrase polypeptides of the present disclosure are used in methods for removing carbon dioxide from a gas stream, *e.g.,* flue gas produced by the combustion of fossil fuels. The methods for removing carbon dioxide from a gas stream *(e.g.,* capturing or extracting CO_2 gas) comprise the step of contacting the gas stream with a solution, wherein comprises

a β-class carbonic anhydrase polypeptide of the disclosure having an improved property *(e.g.,* increased activity, thermostability and/or solvent stability), whereby carbon dioxide from the gas stream is absorbed into the solution (e.g., $C0_2$ gas diffuses into solution and is hydrated to bicarbonate). The method can comprise further steps of isolating and/or separately treating the solution comprising the absorbed carbon dioxide according to known methods to further sequestered and/or otherwise utilize the carbon dioxide. In some embodiments, the methods for removing carbon dioxide from a gas stream can be carried out wherein the recombinant carbonic anhydrase polypeptide is immobilized on a surface, and the method further comprises a step of isolating or separating the immobilized carbonic anhydrase from the solution.

[0020] The methods of removing carbon dioxide from a gas stream using a carbonic anhydrase polypeptide disclosed herein can be carried out under a range of suitable conditions including, but not limited to, polypeptide concentration (and polypeptide form - e.g., lysates, whole cells, or purified powder), solution temperature, solution pH, solution CO_2 loading *(e.g.,* a = 0 to about 0.7), solvent composition, solution concentration of specified CO_2 absorption mediating compound - e.g., an amine compound, ammonia, and/or carbonate ion.

5. DETAILED DESCRIPTION

[0021] The present disclosure is directed to recombinant carbonic anhydrase polypeptides having improved properties, particularly as compared to those of the wild-type carbonic anhydrase polypeptide of *Desulfovibrio vulgaris* having the amino acid sequence of SEQ ID NO: 2. The present disclosure provides the polypeptides, polynucleotides encoding them, vectors and host cells comprising the polynucleotides, and accordingly methods of making the polypeptides including polypeptides that include additional amino acid differences. The present disclosure also provides methods for using such carbonic anhydrases in processes for the capture and sequestration of carbon dioxide *e.g.,* generated by combustion of fossil fuel. The methods disclosed include the use of the recombinant carbonic anhydrase polypeptides under various reaction conditions including conditions comprising the presence of high concentrations of amine co-solvent, ammonia co-solvent, carbonate ions, and/or temperatures that significantly increased or decreased relative to ambient temperatures.

5.1. Definitions

[0022] The technical and scientific terms used in the descriptions herein will have the meanings commonly understood by one of ordinary skill in the art, unless specifically defined otherwise. Accordingly, the following terms are intended to have the following meanings.

[0023] "Carbonic anhydrase" and "CA" are used interchangeably herein to refer to a polypeptide having an enzymatic capability of carrying out the reactions depicted in Scheme 1. Carbonic anhydrase as used herein include naturally occurring (wild-type) carbonic anhydrases as well as non-

naturally occurring, engineered, or recombinant carbonic anhydrase polypeptides generated by human manipulation.

[0024] "Protein", "polypeptide," and "peptide" are used interchangeably herein to denote a polymer of at least two amino acids covalently linked by an amide bond, regardless of length or post-translational modification *(e.g.,* glycosylation, phosphorylation, lipidation, myristilation, ubiquitination, etc.). Included within this definition are D- and L-amino acids, and mixtures of D- and L-amino acids.

[0025] "Naturally occurring" or "wild-type" refers to the form found in nature. For example, a naturally occurring or wild-type polypeptide or polynucleotide sequence is a sequence present in an organism that can be isolated from a source in nature and which has not been intentionally modified by human manipulation.

[0026] "Recombinant" or "engineered" or "non-naturally occurring" when used with reference to, *e.g.,* a cell, nucleic acid, or polypeptide, refers to a material, or a material corresponding to the natural or native form of the material, that has been modified in a manner that would not otherwise exist in nature, or is identical thereto but produced or derived from synthetic materials and/or by manipulation using recombinant techniques. Non-limiting examples include, among others, recombinant cells expressing genes that are not found within the native (non-recombinant) form of the cell or express native genes that are otherwise expressed at a different level.

[0027] "Percentage of sequence identity," "percent identity," and "percent identical" are used herein to refer to comparisons between polynucleotide sequences or polypeptide sequences, and are determined by comparing two optimally aligned sequences over a comparison window, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions *(i.e.,* gaps) as compared to the reference sequence for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which either the identical nucleic acid base or amino acid residue occurs in both sequences or a nucleic acid base or amino acid residue is aligned with a gap to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the window of comparison and multiplying the result by 100 to yield the percentage of sequence identity. Determination of optimal alignment and percent sequence identity is performed using the BLAST and BLAST 2.0 algorithms *(see e.g.,* Altschul et al., 1990, J. Mol. Biol. 215: 403-410 and Altschul et al., 1977, Nucleic Acids Res. 3389-3402). Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information website.

[0028] Briefly, the BLAST analyses involve first identifying high scoring sequence pairs (HSPs) by identifying short words of length W in the query sequence, which either match or satisfy some positive-valued threshold score T when aligned with a word of the same length in a database

sequence. T is referred to as, the neighborhood word score threshold (Altschul et al, supra). These initial neighborhood word hits act as seeds for initiating searches to find longer HSPs containing them. The word hits are then extended in both directions along each sequence for as far as the cumulative alignment score can be increased. Cumulative scores are calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). For amino acid sequences, a scoring matrix is used to calculate the cumulative score. Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T, and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, an expectation (E) of 10, M=5, N=-4, and a comparison of both strands. For amino acid sequences, the BLASTP program uses as defaults a wordlength (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff, 1989, Proc Natl Acad Sci USA 89:10915).

[0029] Numerous other algorithms are available that function similarly to BLAST in providing percent identity for two sequences. Optimal alignment of sequences for comparison can be conducted, *e.g.,* by the local homology algorithm of Smith and Waterman, 1981, Adv. Appl. Math. 2:482, by the homology alignment algorithm of Needleman and Wunsch, 1970, J. Mol. Biol. 48:443, by the search for similarity method of Pearson and Lipman, 1988, Proc. Natl. Acad. Sci. USA 85:2444, by computerized implementations of these algorithms (GAP, BESTFIT, FASTA, and TFASTA in the GCG Wisconsin Software Package), or by visual inspection (see generally, Current Protocols in Molecular Biology, F. M. Ausubel et al., eds., Current Protocols, a joint venture between Greene Publishing Associates, Inc. and John Wiley & Sons, Inc., (1995 Supplement) (Ausubel)). Additionally, determination of sequence alignment and percent sequence identity can employ the BESTFIT or GAP programs in the GCG Wisconsin Software package (Accelrys, Madison WI), using default parameters provided.

[0030] "Reference sequence" refers to a defined sequence to which another sequence is compared. A reference sequence is not limited to wild-type sequences, and can include engineered or altered sequences. For example, a reference sequence can be a previously engineered or altered amino acid sequence. A reference sequence also may be a subset of a larger sequence, for example, a segment of a full-length gene or polypeptide sequence. Generally, a reference sequence is at least 20 nucleotide or amino acid residues in length, at least 25 residues in length, at least 50 residues in length, or the full length of the nucleic acid or polypeptide. Since two polynucleotides or polypeptides may each (1) comprise a sequence *(i.e.,* a portion of the complete sequence) that is similar between the two sequences, and (2) may further comprise a sequence that is divergent between the two sequences,

sequence comparisons between two (or more) polynucleotides or polypeptide are typically performed by comparing sequences of the two polynucleotides over a comparison window to identify and compare local regions of sequence similarity.

[0031] "Comparison window" refers to a conceptual segment of at least about 20 contiguous nucleotide positions or amino acids residues wherein a sequence may be compared to a reference sequence of at least 20 contiguous nucleotides or amino acids and wherein the portion of the sequence in the comparison window may comprise additions or deletions *(i.e.,* gaps) of 20 percent or less as compared to the reference sequence (which does not comprise additions or deletions) for optimal alignment of the two sequences. The comparison window can be longer than 20 contiguous residues, and includes, optionally 30, 40, 50, 100, or longer windows.

[0032] "Corresponding to", "reference to" or "relative to" when used in the context of the numbering of a given amino acid or polynucleotide sequence refers to the numbering of the residues of a specified reference sequence when the given amino acid or polynucleotide sequence is compared to the reference sequence. In other words, the residue number or residue position of a given polymer is designated with respect to the reference sequence rather than by the actual numerical position of the residue within the given amino acid or polynucleotide sequence. For example, a given amino acid sequence, such as that of an engineered carbonic anhydrase, can be aligned to a reference sequence by introducing gaps to optimize residue matches between the two sequences. In these cases, although the gaps are present, the numbering of the residue in the given amino acid or polynucleotide sequence is made with respect to the reference sequence to which it has been aligned.

[0033] "Different from" or "differs from" with respect to a designated reference sequence refers to difference of a given amino acid or polynucleotide sequence when aligned to the reference sequence. Generally, the differences can be determined when the two sequences are optimally aligned. Differences include insertions, deletions, or substitutions of amino acid residues in comparison to the reference sequence.

[0034] "Derived from" as used herein in the context of engineered carbonic anhydrase enzymes, identifies the originating carbonic anhydrase enzyme, and/or the gene encoding such carbonic anhydrase enzyme, upon which the engineering was based.

[0035] "Amino acid residue" or "amino acid" or "residue" as used herein refers to the specific monomer at a sequence position of a polypeptide *(e.g.,* D7 indicates that the "amino acid" or "residue" at position 7 of SEQ ID NO: 2 is an aspartic acid (D).)

[0036] "Amino acid difference" or "residue difference" refers to a change in the amino acid residue at a position of a polypeptide sequence relative to the amino acid residue at a corresponding position in a reference sequence. The positions of amino acid differences generally are referred to herein as "Xn," where n refers to the corresponding position in the reference sequence upon which the residue

difference is based. For example, a "residue difference at position X3 as compared to SEQ ID NO: 2" refers to a change of the amino acid residue at the polypeptide position corresponding to position 3 of SEQ ID NO:2. Thus, if the reference polypeptide of SEQ ID NO: 2 has a glutamine at position 3, then a "residue difference at position X3 as compared to SEQ ID NO:2" an amino acid substitution of any residue other than glutamine at the position of the polypeptide corresponding to position 3 of SEQ ID NO: 2. In most instances herein, the specific amino acid residue difference at a position is indicated as "XnY" where "Xn" specifies the corresponding position as described above, and "Y" is the single letter identifier of the amino acid found in the engineered polypeptide *(i.e.,* the different residue than in the reference polypeptide). In some instances, the present disclosure also provides specific amino acid differences denoted by the conventional notation "AnB", where A is the single letter identifier of the residue in the reference sequence, "n" is the number of the residue position in the reference sequence, and B is the single letter identifier of the residue substitution in the sequence of the engineered polypeptide. For example, "D7S" would refer to the substitution of the amino acid residue, aspartic acid (D) at position 7 of reference sequence with the amino acid serine (S). In some instances, a polypeptide of the present disclosure can include one or more amino acid residue differences relative to a reference sequence, which is indicated by a list of the specified positions where changes are made relative to the reference sequence. The present disclosure includes engineered polypeptide sequences comprising one or more amino acid differences that include either/or both conservative and non-conservative amino acid substitutions.

[0037] "Conservative amino acid substitution" refers to a substitution of a residue with a different residue having a similar side chain, and thus typically involves substitution of the amino acid in the polypeptide with amino acids within the same or similar defined class of amino acids. By way of example and not limitation, an amino acid with an aliphatic side chain may be substituted with another aliphatic amino acid, *e.g.,* alanine, valine, leucine, and isoleucine; an amino acid with hydroxyl side chain is substituted with another amino acid with a hydroxyl side chain, *e.g.,* serine and threonine; an amino acids having aromatic side chains is substituted with another amino acid having an aromatic side chain, *e.g.,* phenylalanine, tyrosine, tryptophan, and histidine; an amino acid with a basic side chain is substituted with another amino acid with a basis side chain, *e.g.,* lysine and arginine; an amino acid with an acidic side chain is substituted with another amino acid with an acidic side chain, *e.g.,* aspartic acid or glutamic acid; and a hydrophobic or hydrophilic amino acid is replaced with another hydrophobic or hydrophilic amino acid, respectively. Exemplary conservative substitutions are provided in Table 1.

Table 1

Residue	Possible Conservative Substitutions
A, L, V, I	Other aliphatic (A, L, V, I) Other non-polar (A, L, V, I, G, M)
G, M	Other non-polar (A, L, V, I, G, M)
D, E	Other acidic (D, E)

K, R	Other basic (K, R)
N, Q, S, T	Other polar
H, Y, W, F	Other aromatic (H, Y, W, F)
C, P	None

[0038] "Non-conservative substitution" refers to substitution of an amino acid in the polypeptide with an amino acid with significantly differing side chain properties. Non-conservative substitutions may use amino acids between, rather than within, the defined groups and affects (a) the structure of the peptide backbone in the area of the substitution *(e.g.,* proline for glycine) (b) the charge or hydrophobicity, or (c) the bulk of the side chain. By way of example and not limitation, an exemplary non-conservative substitution can be an acidic amino acid substituted with a basic or aliphatic amino acid; an aromatic amino acid substituted with a small amino acid; and a hydrophilic amino acid substituted with a hydrophobic amino acid.

[0039] "Deletion" refers to modification of the polypeptide by removal of one or more amino acids from the reference polypeptide. Deletions can comprise removal of 1 or more amino acids, 2 or more amino acids, 5 or more amino acids, 10 or more amino acids, 15 or more amino acids, or 20 or more amino acids, up to 10% of the total number of amino acids, or up to 20% of the total number of amino acids making up the polypeptide while retaining enzymatic activity and/or retaining the improved properties of an engineered carbonic anhydrase enzyme. Deletions can be directed to the internal portions and/or terminal portions of the polypeptide. In various embodiments, the deletion can comprise a continuous segment or can be discontinuous.

[0040] "Insertion" refers to modification of the polypeptide by addition of one or more amino acids to the reference polypeptide. In some embodiments, the improved engineered carbonic anhydrase enzymes comprise insertions of one or more amino acids to the naturally occurring carbonic anhydrase polypeptide as well as insertions of one or more amino acids to other improved carbonic anhydrase polypeptides. Insertions can be in the internal portions of the polypeptide, or to the carboxy or amino terminus. Insertions as used herein include fusion proteins as is known in the art. The insertion can be a contiguous segment of amino acids or separated by one or more of the amino acids in the naturally occurring polypeptide.

[0041] "Fragment" as used herein refers to a polypeptide that has an amino-terminal and/or carboxy-terminal deletion, but where the remaining amino acid sequence is identical to the corresponding positions in the sequence. Fragments can typically have about 80%, 90%, 95%, 98%, and 99% of the full-length carbonic anhydrase polypeptide, for example the polypeptide of SEQ ID NO:2. The amino acid sequences of the specific recombinant carbonic anhydrase polypeptides included in the Sequence Listing of the present disclosure include an initiating methionine (M) residue *(i.e.,* M represents residue position 1). The skilled artisan, however, understands that this initiating methionine residue can be removed by biological processing machinery, such as in a host cell or in vitro translation

system, to generate a mature protein lacking the initiating methionine residue, but otherwise retaining the enzyme's properties. Consequently, the term "amino acid residue difference relative to SEQ ID NO: 2 at position Xn" as used herein may refer to position "Xn" or to the corresponding position *(e.g.,* position (X-l)n) in a reference sequence that has been processed so as to lack the starting methionine.

[0042] "Improved enzyme property" or "improved property" as used herein refers to a functional characteristic of an enzyme that is improved relative to the same functional characteristic of a reference enzyme. Improved enzyme properties of the engineered carbonic anhydrase polypeptides disclosed herein can include but are not limited to: increased thermostability, increased solvent stability, increased pH stability, altered pH activity profile, increased activity (including increased rate conversion of substrate to product, or increased percentage conversion in a period of time), increased and/or altered stereoselectivity, altered substrate specificity and/or preference, decreased substrate, product, and side-product inhibition *(e.g.,* CO_2, carbonate, bicarbonate, carbamate, or solvent-adducts thereof), decreased inhibition by a component of the feedstock *(e.g.* exhaust, flue gas, *etc.),* decreased side-product or impurity production, altered cofactor preference, increased expression, increased secretion, as well as increased stability and/or activity in the presence of additional compounds reagents useful for absorption or sequestration of carbon dioxide, including, for example, amine solvents such as monoethanolamine, methyldiethanolamine, and 2-aminomethylpropanolamine.

[0043] "Stability in the presence of as used in the context of improved enzyme properties disclosed herein refers to stability of the enzyme measured during or after exposure of the enzyme to certain compounds/reagents/ions *(e.g.,* amine compound, ammonia, and/or carbonate ions) in the same solution with the enzyme. It is intended to encompass challenge assays of stability where the enzyme is first exposed to the amine compound or ammonia for some period of time then assayed in a solution under different conditions.

[0044] "Solution" as used herein refers to any medium, phase, or mixture of phases, in which the carbonic anhydrase polypeptide is active. It is intended to include purely liquid phase solutions *(e.g.,* aqueous, or aqueous mixtures with co-solvents, including emulsions and separated liquid phases), as well as slurries and other forms of solutions having mixed liquid-solid phases.

[0045] "Thermostability" refers to the functional characteristic of retaining activity *(e.g.,* more than 60% to 80%) in the presence of, or after exposure to for a period of time *(e.g.* 0.5-24 hrs), elevated temperatures *(e.g.* 30-100°C) compared to the activity of an untreated enzyme.

[0046] "Solvent stability" refers to the functional characteristic of retaining activity *(e.g.,* more than 60% to 80%) in the presence of, or after exposure to for a period of time *(e.g.* 0.5-24 hrs), increased concentrations *(e.g.,* 5-99%) of solvent compared to the activity of an untreated enzyme.

[0047] "pH stability" refers to the functional characteristic of retaining activity *(e.g.,* more than 60% to 80%) in the presence of, or after exposure to for a period of time *(e.g.* 0.5-24 hrs), conditions of high or low pH *(e.g.,* pH 9 to 12) compared to the activity of an untreated enzyme.

[0048] "Increased enzymatic activity" or "increased activity" refers to an improved property of the engineered enzyme *(e.g.,* carbonic anhydrase), which can be represented by an increase in specific activity *(e.g.,* product produced/time/weight protein) or an increase in percent conversion of the substrate to the product *(e.g.,* percent conversion of carbon dioxide to bicarbonate and/or carbonate in a specified time period using a specified amount of carbonic anhydrase) as compared to a reference enzyme under suitable reaction conditions. Exemplary methods to determine enzyme activity are provided in the Examples. Any property relating to enzyme activity may be affected, including the classical enzyme properties of K_m, V_{max} or k_{cat}, changes of which can lead to increased enzymatic activity. Improvements in enzyme activity can be from about 1.1-times the enzymatic activity of the corresponding wild-type carbonic anhydrase enzyme, to as much as 1.2-times, 1.5-times, 2-times, 3-times, 4-times, 5-times, 6-times, 7-times, or more than 8-times the enzymatic activity than the naturally occurring parent carbonic anhydrase. It is understood by the skilled artisan that the activity of any enzyme is diffusion limited such that the catalytic turnover rate cannot exceed the diffusion rate of the substrate, including any required cofactors. The theoretical maximum of the diffusion limit, or k_{cat}/K_m, is generally about 10^8 to 10^9 (M^{-1} s^{-1}). Hence, any improvements in the enzyme activity of the carbonic anhydrase will have an upper limit related to the diffusion rate of the substrates acted on by the carbonic anhydrase enzyme. Carbonic anhydrase activity can be measured by any one of standard assays used for measuring carbonic anhydrase, *e.g.,* as provided in the Examples. Comparisons of enzyme activities are made, *e.g.,* using a defined preparation of enzyme, a defined assay under a set of conditions, as further described in detail herein. Generally, when lysates are compared, the numbers of cells and the amount of protein assayed are determined as well as use of identical expression systems and identical host cells to minimize variations in amount of enzyme produced by the host cells and present in the lysates.

[0049] "Conversion" refers to the enzymatic conversion of the substrate to the corresponding product. "Percent conversion" refers to the percent of the substrate that is reduced to the product within a period of time under specified conditions. Thus, the "enzymatic activity" or "activity" of a carbonic anhydrase polypeptide can be expressed as "percent conversion" of the substrate to the product.

[0050] "Isolated polypeptide" refers to a polypeptide which is substantially separated from other contaminants that naturally accompany it, *e.g.,* protein, lipids, and polynucleotides. The term embraces polypeptides which have been removed or purified from their naturally-occurring environment or expression system *(e.g.,* host cell or in vitro synthesis). The improved carbonic anhydrase enzymes may be present within a cell, present in the cellular medium, or prepared in

various forms, such as lysates or isolated preparations. As such, in some embodiments, the improved carbonic anhydrase enzyme can be an isolated polypeptide.

[0051] "Substantially pure polypeptide" refers to a composition in which the polypeptide species is the predominant species present *(i.e.,* on a molar or weight basis it is more abundant than any other individual macromolecular species in the composition), and is generally a substantially purified composition when the object species comprises at least about 50 percent of the macromolecular species present by mole or % weight. Generally, a substantially pure carbonic anhydrase composition will comprise about 60 % or more, about 70% or more, about 80% or more, about 90% or more, about 95% or more, and about 98% or more of all macromolecular species by mole or % weight present in the composition. In some embodiments, the object species is purified to essential homogeneity *(i.e.,* contaminant species cannot be detected in the composition by conventional detection methods) wherein the composition consists essentially of a single macromolecular species. Solvent species, small molecules (<500 Daltons), and elemental ion species are not considered macromolecular species. In some embodiments, the isolated improved carbonic anhydrase polypeptide is a substantially pure polypeptide composition.

[0052] "Coding sequence" refers to that portion of a polynucleotide that encodes an amino acid sequence of a protein *(e.g.,* a gene).

[0053] "Heterologous" polynucleotide refers to any polynucleotide that is introduced into a host cell by laboratory techniques, and includes polynucleotides that are removed from a host cell, subjected to laboratory manipulation, and then reintroduced into a host cell.

[0054] "Codon optimized" refers to changes in the codons of the polynucleotide encoding a protein to those preferentially used in a particular organism such that the encoded protein is efficiently expressed in the organism of interest. In some embodiments, the polynucleotides encoding the carbonic anhydrase enzymes may be codon optimized for optimal production from the host organism selected for expression.

[0055] "Control sequence" is defined herein to include all components, which are necessary or advantageous for the expression of a polynucleotide and/or polypeptide of the present disclosure. Each control sequence may be native or foreign to the polynucleotide of interest. Such control sequences include, but are not limited to, a leader, polyadenylation sequence, propeptide sequence, promoter, signal peptide sequence, and transcription terminator.

[0056] "Operably linked" is defined herein as a configuration in which a control sequence is appropriately placed *(i.e.,* in a functional relationship) at a position relative to a polynucleotide of interest such that the control sequence directs or regulates the expression of the polynucleotide and/or polypeptide of interest.

5.2. Carbonic Anhydrase Polypeptides

[0057] The recombinant (or engineered) carbonic anhydrase ("CA") polypeptides of the present disclosure are those having an improved property when compared with a naturally-occurring, wild type carbonic anhydrase enzyme obtained from *Desulfovibrio vulgaris* (SEQ ID NO: 2). Enzyme properties for which improvement is desirable include, but are not limited to, activity *(e.g.,* hydration of carbon dioxide, or dehydration of bicarbonate), thermal stability, solvent stability, pH activity profile, refractoriness to inhibition or inactivation by other compounds in the solution with the enzyme, *e.g.* inhibition by bicarbonate, carbonate, amine compounds, ammonia, flue gas components, etc.. The improvements can relate to a single enzyme property, such as activity, or a combination of properties, such as activity and solvent stability and/or thermostability.

[0058] In some embodiments, the carbonic anhydrase polypeptides provided herein comprise one or more differences in their amino acid sequence *(e.g.,* substitutions, insertions, and/or deletions) relative to a reference sequence *(e.g., Desulfovibrio vulgaris* CA polypeptide of SEQ ID NO: 2) that result in a carbonic anhydrase polypeptide having an improved property. In such embodiments, the number of differences in the amino acid sequence can comprise differences in 1 or more amino acids, 2 or more amino acids, 3 or more amino acids, 4 or more amino acids, 5 or more amino acids, 6 or more amino acids, 8 or more amino acids, 9 or more amino acids, 10 or more amino acids, 15 or more amino acids, or 20 or more amino acids, up to 10% of the total number of amino acids, up to 10% of the total number of amino acids, up to 20% of the total number of amino acids, or up to 30% of the total number of amino acids of the reference enzyme sequence. In some embodiments, the number of differences in the carbonic anhydrase polypeptide that produces an improved carbonic anhydrase property may comprise from about 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-25, 1-30, 1-35 or about 1-40 differences relative the reference sequence. The differences can comprise insertions, deletions, substitutions, or combinations thereof.

[0059] In some embodiments, the improved property of the carbonic anhydrase polypeptide is with respect to an increase in its ability to convert a greater percentage of the substrate to the product. In some embodiments, the improved property of the carbonic anhydrase polypeptide is with respect to an increase in its rate of conversion of the substrate to the product *(e.g.,* hydration of carbon dioxide to bicarbonate). This improvement in enzymatic activity can be manifested by the ability to use less of the improved polypeptide as compared to the wild-type or other reference sequence(s) to reduce or convert the same amount of product. In some embodiments, the improved property of the carbonic anhydrase polypeptide is with respect to its thermostability. Accordingly, in some embodiments the carbonic anhydrase polypeptides have an improved property that comprises at least 1.2-fold, at least 1.3-fold, at least 1.5-fold, at least 2-fold, at least 3-fold, at least 4-fold, at least 5-fold, at least 10-fold, or at least 25-fold increased thermostability. In such embodiments, increased thermostability can be determined as increased activity relative to a reference polypeptide following exposure to

thermostability challenge conditions - e.g., exposure to 30, 40, 50, or 60°C solution for a defined time period, such as 24 h. In some embodiments, the carbonic anhydrase polypeptide has more than one improved property, such as a combination of increased enzyme activity and thermostability.

[0060] In some embodiments, the improved property of the carbonic anhydrase polypeptides disclosed herein is increased stability in the presence of compounds in the enzyme solution that improve the ability of the solution to absorb carbon dioxide *(i.e.,* compounds that mediate the absorption of $C0_2$ by the solution). Such compounds increase the amount of carbon dioxide that the solution can absorb, increase the rate at which carbon dioxide is absorbed, and/or improve the thermodynamic properties of the solution that control the carbon dioxide absorption or desorption. Accordingly, the carbonic anhydrases disclosed herein are advantageous for use in methods for carbon dioxide capture and sequestration that use solutions into which carbon dioxide is absorbed *(i.e.,* captured by diffusing from gas stream into the liquid solution) and/or from which carbon dioxide is desorbed *(i.e.,* extracted by diffusing from liquid solution into gas phase). Such compounds, solutions, and solvent systems for the absorption and/or desorption of carbon dioxide and the associated processes of using them for carbon dioxide capture from gas streams are described in *e.g.,* U.S. Pat. Nos. 6,143,556, 6,524,843 B2, 7,176,017 B2, 7,596,952 B2, 7,641,717 B2, 7,579185 B2, 7,740,689 B2, 7,132,090 B2; U.S. Pat. Publ. Nos. 2007/0256559A1, 2009/0 155889A1, 2010/0086983A1; PCT Publ. Nos. WO2006/089423A1, WO2008/072979A1, WO2009/000025A1, WO201 0/0200 17A1, WO20 10/01 4773 Al, WO2010/045689A1, each of which is hereby incorporated by reference herein.

[0061] In some embodiments, the improved property of the carbonic anhydrase polypeptide is increased stability in the presence of an amine compound in the enzyme solution. In addition to increased stability to the presence of amine compound, in such embodiments the carbonic anhydrase can have increased thermostability, *e.g.,* increased activity at temperatures above 40°C. The carbonic anhydrases disclosed herein having increased stability to amine compounds and increased solution temperature are particularly advantageous for use in methods for carbon dioxide capture and sequestration from flue gas streams using solutions comprising amine compounds (see *e.g.,* U.S. Pat. No. 7,740,689 B2, and U.S. Pat. Publ. 2009/0155889 Al, each of which is hereby incorporated by reference herein), such as those amine compounds selected from: 2-(2-aminoethylamino)ethanol (AEE), 2-amino-2-hydroxymethyl- 1,3-propanediol (AHPD), 2-amino-2-methyl-l-propanol (AMP), diethanolamine (DEA), diisopropanolamine (DIPA), N-hydroxyethylpiperazine (HEP), N-methyldiethanolamine (MDEA), monoethanolamine (MEA), N-methylpiperazine (MP), piperazine, piperidine, 2-(2-tert-butylaminoethoxy)ethanol (TBEE), triethanolamine (TEA), triisopropanolamine (TIA), tris, 2-(2-aminoethoxy)ethanol, 2-(2-tert-butylaminopropoxy)ethanol, 2-(2-tert-amylaminoethoxy)ethanol, 2-(2-isopropylaminopropoxy)ethanol, 2-(2-(l -methyl- 1-ethylpropylamino)ethoxy)ethanol, and mixtures thereof.

[0062] In some embodiments, the improved property of the carbonic anhydrase polypeptide is increased stability in the presence of ammonia in the enzyme solution. In addition to increased stability to the presence of ammonia, in such embodiments the carbonic anhydrase can have increased stability at increased or decreased temperatures *(e.g.,* less than about 15°C). The carbonic anhydrases disclosed herein having increased stability to ammonia and/or increased thermostability are particularly advantageous for use in methods for carbon dioxide capture and sequestration from flue gas streams using solutions comprising ammonia, such as the chilled ammonia processes (see *e.g.,* U.S. Pat. No. 7,641,717 B2, U.S. Pat. Publ. 2009/0155889 Al, each of which is hereby incorporated by reference herein).

[0063] Exemplary recombinant carbonic anhydrase polypeptides of the present disclosure having an improved enzyme property relative to the wild-type *Desulfovibrio vulgaris* reference polypeptide of SEQ ID NO:2 include but are not limited to, the polypeptides that comprise the amino acid sequences corresponding to any one of SEQ ID NOs: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 240, 242, 244, 246, 248, 250, 252, 254, 256, 258, 260, 262, 264, 266, and 268.

[0064] Structure and function information for exemplary recombinant carbonic anhydrase polypeptides of the present disclosure are shown below in Tables 2A and 2B. The odd numbered sequence identifiers *(i.e.,* SEQ ID NOs) refer to the nucleotide sequence encoding the amino acid sequence provided by the even numbered SEQ ID NOs, and the sequences are provided in the electronic sequence listing file accompanying this disclosure, which is hereby incorporated by reference herein. The amino acid residue differences are based on comparison to the reference sequence of SEQ ID NO: 2, which is a wild type carbonic anhydrase from *Desulfovibrio vulgaris* str. "Miyazaki F" having GenBank accession ACL09337.1 GL218758438.

[0065] Tables 2A and 2B also disclose the increased stability (solvent and/or thermostability) in the presence of an amine compound (MDEA) and/or ammonia at various concentrations and temperatures relative to the reference polypeptide of SEQ ID NO: 2. Increased stability was determined by measuring the relative rate of dehydrating bicarbonate to carbon dioxide in a high-throughput (HTP) assay following 24 h exposure to the specified solvent and temperature challenge conditions. HTP activity assays were carried out in 96-well plate format assay using cell lysates containing the engineered polypeptides. General HTP challenge/assay conditions were as follows: 25 µï¿½ of cleared *E. coli* lysate added to 75 µï¿½ of challenge buffer solution *(e.g.,* solution containing 4.0 M - 6.66 M MDEA or NH_3) and incubated at the challenge temperature *(e.g.,* 30°, 35°, 42°, 50° or 55°C) for 24 h;

followed by adding a 10 μL aliquot of the challenge solution to 190 μL of bicarbonate dehydration assay solution (200 mM $KHC0_3$, 400 μM phenolphthalem, pH 7 or 8) at 25°C or 45°C, measuring carbonic anhydrase activity as slope of phenolphthalem indicator absorbance change at 550 nm over 20-30 minutes. Additional HTP assay details are provided in Example 1. As noted in Tables 2A and 2B, the measured level of increased activity of each engineered polypeptide relative to a reference polypeptide was classified as "+", "++", or "+++" for the different assays.

TABLE 2A

SEQ ID NO: (nt/aa)	**Amino Acid Residue Difference(s)** (relative to SEQ ID NO: 2)	**Assay 1** (24h/42°C/ 3M MDEA challenge/ 25°C assay)	**Assay 2** (24h/50°C/ 3M MDEA challenge/ 25°C assay)	**Assay 3** (24h/30°C/ 4.2M NH_3/CO_2 challenge/ 25°C assay)	**Assay 4** (24h/35°C/ 4.2M NH_3/CO_2 challenge/ 25°C assay)
3/4	K147E;	+++	+++		
5/6	T30R;	+++		++	+++
7/8	T139M;	+++	++		
9/10	G120R;	+++			
11/12	T30Q;	+++		++	++
13/14	T4F;	+++	+	+	++
15/16	A84Q;	+++	+++	++	++
17/18	Q119M;	+++			
19/20	L34H;	+++		++	++
21/22	T4M; T30K;	++		+++	+++
23/24	K147T;	++	+++		
25/26	A56S;	++	+++	++	++
27/28	Q32K;	++		+	++
29/30	V131L;	++	+	+	
31/32	Q15R; T30R;	++	++	+++	+++
33/34	N145W;	++			
35/36	R16S;	++			
37/38	A40W;	++		+	+
39/40	N213E;	++			
41/42	H222C;	++		++	+++
43/44	E142L;	++			
45/46	G2T;	++	++		
47/48	R31P;	++	+		
49/50	S144L;	++			
51/52	E159H;	++			
53/54	T139Q;	++			
55/56	H148T;	++			
57/58	M170F;	++	+	+	
59/60	D86A				
61/62	A121K;	++	+	+	
63/64	N145F;	++			

65/66	Q32R;	++		++	+
67/68	A121W;	++	+	+	++
69/70	K37R;	++		+	++
71/72	A221C;	++		++	++
73/74	A84S;	++			
75/76	E200R;	++			
77/78	T139K;	++			
79/80	A95V;	++		+	+
81/82	A84N;	+		+	
83/84	Q43M;	+			
85/86	A121V;	+			
87/88	K147G;	+			
89/90	R223C;	+		++	++
91/92	T30A;	+		++	++
93/94	G2R;	+		+	+
95/96	A121H;	+		+	
97/98	A121Q;	+		+	+
99/100	A60C;	+		++	++
101/102	D96C;	+			
103/104	T30L;	+		++	+
105/106	A40L;	+	+	+	
107/108	H97F;	+			
109/110	E68A;	+		+	+
111/112	S42A; A219T;	+		+	+
113/114	V70I;	+		++	++
115/116	Q119T;	+			
117/118	D96E;	+		+	
119/120	S35A;	+		+	++
121/122	H124G;	+		+	
123/124	Q119K;	+			
125/126	V138L;	+			
127/128	D168E;	+			
129/130	T139H;	+			
131/132	A121T;	+		+	+
133/134	A121L;	+			
135/136	S144A;	+			
137/138	N145C;	+			
139/140	N213Q;	+			
141/142	D96K;	+			
143/144	A178G;	+			
145/146	H124R;	+			
147/148	D96A;	+			
149/150	S35R;	+		+	+
151/152	E159V;	+			

153/154	T47R;	+		+	
155/156	H148A;	+			
157/158	A84R;	+		++	++
159/160	Q43V;	+			
161/162	E159R;	+			
163/164	K147F;	+			
165/166	E68G;	+		+	+
167/168	V157A;			+	+
169/170	V138W;			+	
171/172	V138F;			+	
173/174	R223Q;			+	++
175/176	M207E;			+	
177/178	A84K;			++	+
179/180	A60V;			++	++
181/182	A40Q;			++	+
183/184	A22G;			++	
185/186	K143M; M207N;				

Levels of increased activity were determined relative to the reference polypeptide of SEQ ID NO: 2 and defined as"+", "++", or "+++" for each of the four assays as follows:

Assay 1: "+" = at least 1.3-fold but less than 2-fold increased activity; "++"= at least 2-fold but less than 3-fold increased activity; "+++" = at least 3-fold increased activity.

Assay 2: "+" = at least 1.5-fold but less than 2-fold increased activity; "++" = at least 2-fold but less than 3-fold increased activity; "+++" = at least 3-fold increased activity.

Assay 3: "+" = at least 1.3-fold but less than 1.5-fold increased activity; "++" = at least 1.5-fold but less than 2-fold increased activity; "+++" = at least 2-fold increased activity.

Assay 4: "+" = at least 1.3-fold but less than 3-fold increased activity; "++" = at least 3-fold but less than 5-fold increased activity; "+++" = at least 5-fold increased activity.

TABLE 2B

SEQ ID NO: (nt/aa)	**Amino Acid Residue Difference(s)** (relative to SEQ ID NO: 2)	**Assay 5** (24h/50°C/ 4M MDEA challenge/ 45°C assay)	**Assay 6** (24h/50°C/ 5M MDEA challenge/ 25°C assay)	**Assay 7** (24h/55°C/ 5M MDEA challenge/ 25°C assay)
187/188	T30R; R31P; A56S; A84Q;	++	+++	+++
189/190	A56S; A84Q; T139M;	++	+++	+++
191/192	T30R; R31P; A40L; A56S; G120R;	+++	+++	+++
193/194	R31P; A40L; A56S; G120R; T139M;	+++	+++	+++
195/196	T30R; R31P; A56S; A84Q; Q119M;	+++	+++	+++

197/198	R31P; A40L; A56S; A84Q;	+++	+++	+++
199/200	T30Q; R31P; A56S; A84Q;	+++	+++	+++
201/202	T30Q; L34H; A56S;	+++	+++	++
203/204	T30R; R31P; A40L; A56S; K147T;	+++	+++	+++
205/206	T30R; R31P; A56S; K147T;	+++	+++	+++
207/208	T4F; A56S; A84Q;	++	++	++
209/210	T30R; L34H; A56S;	++	++	++
211/212	A56S; T139M;	+	+	+
213/214	G2T; R31P; L34H; A40L; A56S; A84Q; T139M;	+	+	++
215/216	T4F; L34H; A56S; G120R; K147E;	+	+	+
217/218	A40L; A56S;	++	+	+
219/220	R31P; A40L; A56S; Q119M; G120R;	++	+	++
221/222	R31P; A56S; G120R; K147E;	++	+	
223/224	T4F; A40L; A56S; K147T;	++	+	+
225/226	R31P; A40L; A56S;	++	+	+
227/228	A56S; A84Q;	++	+	+
229/230	T30R; A40L; A56S;	++	+	++
231/232	T30Q; L34H; A56S; K147T;	++	+	++
233/234	L34H; A56S;	++	+	+
235/236	T30R; R31P; A56S;	+++	+	++
237/238	T30R; A56S;	+++	+	+
239/240	R31P; A56S; A84Q;	+++	+	++
241/242	T4F; A56S;			+
243/244	G2T; A56S; T139M;			+
245/246	A56S; G120R; K147T;	+		
247/248	G2T; A56S; A84Q; T139M;	+		+
249/250	A56S; Q119L; G120R;	+		+
251/252	A40L; A56S; G120R;	+		+
253/254	A56S; K147T;	+		
255/256	A40L; A56S; T139M; K147E;	+		
257/258	A40L; A56S; T139M;	++		
259/260	T4F; T30Q; A56S; G120R; T139M;	++		+
261/262	L34H; A56S; A84Q; T139M;	++		++
263/264	A56S; A84Q; K147E;	++		+
265/266	A56S; A84Q; G120R;	++		+
267/268	T30R; R31P; A56S; T139M;	++		++

Levels of increased activity were determined relative to the reference polypeptide of SEQ ID NO: 26 (*i.e.*, engineered polypeptide having A56S) and defined as"+", "++", or "+++" for each of the four assays as follows: Assay 5: "+" indicates at least 1.5-fold but less than 2.5-fold increased activity; "++" indicates at least 2.5-fold

but less than 4-fold increased activity; "+++"indicates at least 4-fold increased activity.
Assay 6: "+" indicates at least 1.3-fold but less than 1.7-fold increased activity; "++" indicates at least 1.7-fold but less than 2-fold increased activity; "+++" indicates at least 2-fold increased activity.
Assay 7: "+" indicates at least 1.5-fold but less than 2.5-fold increased activity; "++" indicates at least 2.5-fold but less than 4-fold increased activity; "+++"indicates at least 4-fold increased activity.

[0066] In addition to the exemplary polypeptides of Tables 2A and 2B, in some embodiments, the present disclosure provides a recombinant carbonic anhydrase polypeptide having an improved enzyme property relative to a polypeptide of SEQ ID NO:2, and an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to a reference amino acid sequence selected from any one of SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 240, 242, 244, 246, 248, 250, 252, 254, 256, 258, 260, 262, 264, 266, and 268.

[0067] Each of the exemplary recombinant carbonic anhydrase polypeptides shown in Tables 2A and 2B comprises one or more amino acid residue differences as compared to SEQ ID NO: 2, and has at least 1.3-fold, at least 1.5-fold, at least 2-fold, at least 3-fold, or at least 5-fold increased stability relative to the polypeptide of SEQ ID NO: 2. Specific amino acid differences are shown in Tables 2A and 2B and include one or more residue differences as compared to SEQ ID NO:2 at the following residue positions: X2; X4; X15; X16; X22; X30; X31; X32; X34; X35; X37; X40; X42; X43; X47; X56; X60; X68; X70; X84; X86; X95; X96; X97; X I 19; X120; X121; X124; X131; X138; X139; X142; X143; X144; X145; X147; X148; X157; X159; X168; X170; X178; X200; X207; X213; X219; X221; X222; and X223. Some of these positions appear in more than one polypeptide with different amino acid replacements. The specific amino acid residue differences found in the exemplary polypeptides having an improved property are: X2R; X2T; X4F; X4M; X I5R; X I6S; X22G; X30A; X30K; X30L; X30Q; X30R; X31P; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40Q; X40W; X42A; X43M; X43V; X47R; X56S; X60C; X60V; X68A; X68G; X70I; X84K; X84N; X84Q; X84R; X84S; X86A; X95V; X96A; X96C; X96E; X96K; X97F; X I 19K; X I 19L; X I 19M; X I 19T; X120R; X121H; X121K; X121L; X121Q; X121T; X121V; X121W; X124G; X124R; X131L; X138F; X138L; X138W; X139H; X139K; X139M; X139Q; X142L; X143M; X144A; X144L; X145C; X145F; X145W; X147E; X147F; X147G; X147T; X148A; X148T; X157A; X159H; X159R; X159V; X168E; X170F; X178G; X200R; X207E; X207N; X213E; X213Q; X219T; X221C; X222C; X223C; and X223Q.

[0068] It will be apparent to the skilled artisan that the residue positions and specific residue differences of the present disclosure which have been shown to improve stability in solutions comprising amine compounds and/or ammonia can be used to generate recombinant carbonic anhydrase polypeptides besides the exemplary polypeptides of Tables 2A and 2B. It is contemplated that additional recombinant carbonic anhydrase polypeptides having improved properties can be prepared comprising various combinations of the amino acid residue differences of the exemplary polypeptides of Tables 2A and 2B. This has been demonstrated by the recombinant carbonic anhydrase polypeptides of Table 2B which were prepared by combining the single amino acid difference of SEQ ID NO: 26 *(i.e.,* X56S) with various other amino acid differences from Table 2A to create the improved polypeptides of SEQ ID NO: 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 240, 242, 244, 246, 248, 250, 252, 254, 256, 258, 260, 262, 264, 266, and 268.

[0069] Accordingly, in some embodiments, the present disclosure provides a recombinant carbonic anhydrase polypeptide having at least 1.3-fold, at least 1.5-fold, at least 2-fold, at least 3-fold, or at least 5-fold increased stability relative to the polypeptide of SEQ ID NO: 2, comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 2, and further comprises the one or more amino acid residue differences as compared to SEQ ID NO:2 of any one of SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 240, 242, 244, 246, 248, 250, 252, 254, 256, 258, 260, 262, 264, 266, and 268. In some embodiments, in addition to the set of amino acid residue differences of any one of the recombinant carbonic anhydrase polypeptides of SEQ ID NO: 4 through SEQ ID NO: 186, the sequence of the recombinant polypeptide can further comprise 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-11, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-26, 1-30, 1-35, 1-40 residue differences at other amino acid residue positions as compared to the SEQ ID NO: 2. In some embodiments, the residue differences can comprise conservative substitutions and/or non-conservative substitutions as compared to SEQ ID NO: 2.

[0070] In some embodiments, any of the recombinant carbonic anhydrase polypeptides having an improved property relative to the polypeptide of SEQ ID NO: 2 and an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 2 and one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2R; X2T; X4F; X4M; X15R; X16S; X22G; X30A; X30K; X30L; X30Q; X30R;

X31P; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40Q; X40W; X42A; X43M; X43V; X47R; X56S; X60C; X60V; X68A; X68G; X70I; X84K; X84N; X84Q; X84R; X84S; X86A; X95V; X96A; X96C; X96E; X96K; X97F; X I 19K; X I 19L; X I 19M; X I 19T; X120R; X121H; X121K; X121L; X121Q; X121T; X121V; X121W; X124G; X124R; X131L; X138F; X138L; X138W; X139H; X139K; X139M; X139Q; X142L; X143M; X144A; X144L; X145C; X145F; X145W; X147E; X147F; X147G; X147T; X148A; X148T; X157A; X159H; X159R; X159V; X168E; X170F; X178G; X200R; X207E; X207N; X213E; X213Q; X219T; X221C; X222C; X223C; and X223Q.

[0071] The positions associated with the improved property of increased stability in the presence of an amine compound include: X2; X4; X I5; X I6; X30; X31; X32; X34; X35; X37; X40; X42; X43; X47; X56; X60; X68; X70; X84; X86; X95; X96; X97; X I 19; X120; X121; X124; X131; X138; X139; X142; X144; X145; X147; X148; X159; X168; X170; X178; X200; X213; X219; X221; X222; and X223. The specific amino acid residue differences associated with the improved property of increased stability in the presence of an amine compound include: X2R; X2T; X4F; X4M; X15R; X I6S; X30A; X30K; X30L; X30Q; X30R; X31P; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40W; X42A; X43M; X43V; X47R; X56S; X60C; X68A; X68G; X70I; X84N; X84Q; X84R; X84S; X86A; X95V; X96A; X96C; X96E; X96K; X97F; X I 19K; X I 19L; X I 19M; X I 19T; X120R; X121H; X121K; X121L; X121Q; X121T; X121V; X121W; X124G; X124R; X131L; X138L; X139H; X139K; X139M; X139Q; X142L; X144A; X144L; X145C; X145F; X145W; X147E; X147F; X147G; X147T; X148A; X148T; X159H; X159R; X159V; X168E; X170F; X178G; X200R; X213E; X213Q; X219T; X221C; X222C; andX223C.

[0072] In some embodiments, the present disclosure provides a recombinant carbonic anhydrase polypeptide in which the increased stability in the presence of an amine compound in which the increased stability in the presence of an amine compound comprises at least 1.3-fold increased activity relative to the reference polypeptide of SEQ ID NO: 2 after 24 hours exposure to 4 M MDEA at 42°C and the amino acid sequence comprises one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2R; X2T; X4F; X4M; X15R; X16S; X30A; X30K; X30L; X30Q; X30R; X31P; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40W; X42A; X43M; X43V; X47R; X56S; X60C; X68A; X68G; X70I; X84N; X84Q; X84R; X84S; X86A; X95V; X96A; X96C; X96E; X96K; X97F; X I 19K; X I 19L; X I 19M; X I 19T; X120R; X121H; X121K; X121L; X121Q; X121T; X121V; X121W; X124G; X124R; X131L; X138L; X139H; X139K; X139M; X139Q; X142L; X144A; X144L; X145C; X145F; X145W; X147E; X147F; X147G; X147T; X148A; X148T; X159H; X159R; X I59V; X168E; X170F; X178G; X200R; X213E; X213Q; X219T; X221C; X222C; and X223C.

[0073] In some embodiments, the present disclosure provides a recombinant carbonic anhydrase polypeptide in which the increased stability in the presence of an amine compound in which the increased stability in the presence of an amine compound comprises at least 1.5-fold increased activity

relative to the reference polypeptide of SEQ ID NO: 2 after 24 hours exposure to 4 M MDEA at 50°C and the amino acid sequence comprises one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2T; X4F; X31P; X40L; X56S; X84Q; X119M; X120R; X121K; X121W; X131L; X139M; X147E; X147T; and X170F.

[0074] In some embodiments, the present disclosure provides a recombinant carbonic anhydrase polypeptide in which the increased stability in the presence of an amine compound in which the increased stability in the presence of an amine compound comprises at least 2-fold increased activity relative to the reference polypeptide of SEQ ID NO: 2 after 24 hours exposure to 4 M MDEA at 50°C and the amino acid sequence comprises one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2T; X56S; X84Q; X139M; X147E; and X147T.

[0075] In some embodiments, the present disclosure provides a recombinant carbonic anhydrase polypeptide in which the increased stability in the presence of an amine compound in which the increased stability in the presence of an amine compound comprises at least 3-fold increased activity relative to the reference polypeptide of SEQ ID NO: 2 after 24 hours exposure to 4 M MDEA at 50°C and an amino acid sequence comprising one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X56S; X84Q; X147E and X147T.

[0076] In some embodiments, the present disclosure provides a recombinant carbonic anhydrase polypeptide in which the increased stability in the presence of an amine compound in which the amino acid sequence comprises the amino acid difference X56S and one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2R; X2T; X4F; X4M; X15R; X16S; X30A; X30K; X30L; X30Q; X30R; X31P; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40W; X42A; X43M; X43V; X47R; X60C; X68A; X68G; X70I; X84N; X84Q; X84R; X84S; X86A; X95V; X96A; X96C; X96E; X96K; X97F; X119K; X119L; X119M; X119T; X120R; X121H; X121K; X121L; X121Q; X121T; X121V; X121W; X124G; X124R; X131L; X138L; X139H; X139K; X139M; X139Q; X142L; X144A; X144L; X145C; X145F; X145W; X147E; X147F; X147G; X147T; X148A; X148T; X159H; X159R; X159V; X168E; X170F; X178G; X200R; X213E; X213Q; X219T; X221C; X222C; and X223C.

[0077] The positions associated with the improved property of increased stability in the presence of an ammonia include: X2; X4; X15; X22; X30; X32; X34; X35; X37; X40; X42; X47; X56; X60; X68; X70; X84; X86; X95; X96; X121; X124; X138; X143; X157; X170; X207; X219; X221; X222; and X223. The specific amino acid residue differences associated with the improved property of increased stability in the presence of ammonia include: X2R; X4F; X4M; X15R; X22G; X30A; X30K; X30L; X30Q; X30R; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40Q; X40W; X42A; X47R; X56S; X60C; X60V; X68A; X68G; X70I; X84K; X84Q; X86A; X95V; X96E;

X121H; X121K; X121Q; X121T; X121W; X124G; X138F; X138W; X143M; X157A; X170F; X207E; X207N; X219T; X221C; X222C; andX223Q.

[0078] In some embodiments, the present disclosure provides a recombinant carbonic anhydrase polypeptide in which the increased stability in the presence of ammonia comprises at least 1.3-fold increased activity relative to the reference polypeptide of SEQ ID NO: 2 after 24 hours exposure to ammonia at 35°C and the amino acid sequence comprises one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2R; X4F; X22G; X30A; X30L; X30Q; X3 OR; X32K; X32R; X34H; X35A; X35R; X37R; X40Q; X40W; X56S; X60C; X60V; X68A; X68G; X70I; X84K; X84Q; X86A; X95V; X121Q; X121T; X121W; X157A; X221C; X222C; andX223Q.

[0079] In some embodiments, the present disclosure provides a recombinant carbonic anhydrase polypeptide in which the increased stability in the presence of ammonia comprises at least 3-fold increased activity relative to the reference polypeptide of SEQ ID NO: 2 after 24 hours exposure to ammonia at 35°C and the amino acid sequence comprises one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X4F; X22G; X30A; X30Q; X30R; X32K; X34H; X35A; X37R; X56S; X60C; X60V; X70I; X84Q; X121W; X221C; X222C; andX223Q.

[0080] In some embodiments, the present disclosure provides a recombinant carbonic anhydrase polypeptide having increased stability in the presence of ammonia in which the amino acid sequence comprises one or more of the amino acid residue differences selected from X I5R and X30R, and further comprises one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2R; X4F; X4M; X I5R; X22G; X30A; X30K; X30L; X30Q; X30R; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40Q; X40W; X42A; X47R; X56S; X60C; X60V; X68A; X68G; X70I; X84K; X84Q; X95V; X96E; X121H; X121K; X121Q; X121T; X121W; X124G; X138F; X138W; X143M; X157A; X170F; X207E; X207N; X219T; X221C; X222C; and X223Q. In some embodiments, the amino acid sequence comprises both X15R and X30R and further comprises one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2R; X4F; X4M; X22G; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40Q; X40W; X42A; X47R; X56S; X60C; X60V; X68A; X68G; X70I; X84K; X84Q; X95V; X96E; X121H; X121K; X121Q; X121T; X121W; X124G; X138F; X138W; X143M; X157A; X170F; X207E; X207N; X219T; X221C; X222C; andX223Q.

[0081] As described in Tables 2A and 2B and the Examples, the improved property of increased stability and/or increased activity are determined under suitable conditions. In some embodiments, improved property comprises at least 1.2-fold, at least 1.3-fold, at least 1.5-fold, at least 2-fold, at least 3-fold, at least 4-fold, at least 5-fold, at least 10-fold, or at least 25-fold increased activity of hydrating carbon dioxide or dehydrating bicarbonate under suitable conditions. In some embodiments, the suitable conditions comprise a carbonic anhydrase polypeptide concentration of from about 0.1 g/L to about 10 g/L, about 0.25 g/L to about 7.5 g/L, about 0.5 g/L to about 5 g/L, less

than 10 g/L, less than about 5 g/L, or less than about 2.5 g/L. In some embodiments, the suitable conditions comprise a loading of solution CO_2 of from about a = 0.005 to about a = 0.4, from about a = 0.01 to about a = 0.3, a = 0.015 to about a = 0.25, a = 0.02 to about a = 0.2, less than about a = 0.3, less than about a = 0.25, or less than about a = 0.2.

[0082] In some embodiments the improved property is activity measured after exposure of the carbonic anhydrase to thermal or solvent challenge conditions. Accordingly in some embodiments, the increased activity is determined after heating the recombinant carbonic anhydrase polypeptide and the reference polypeptide at a temperature of from about 30°C to 60°C for a period of time of about 60 minutes to about 1440 minutes. In such embodiments, the fold-increase in activity corresponds to the same fold-increase in thermostability or solvent stability - depending on the challenge conditions. Various other challenge conditions may be used as disclosed in the Examples and elsewhere herein.

[0083] In some embodiments the improved property is stability in the presence of an amine compound and the suitable conditions comprise a solution comprising an amine compound selected from: 2-(2-aminoethylamino)ethanol (AEE), 2-amino-2-hydroxymethyl- 1,3-propanediol (AHPD), 2-amino-2-methyl- 1-propanol (AMP), diethanolamine (DEA), diisopropanolamine (DIPA), N-hydroxyethylpiperazine (HEP), N-methyldiethanolamine (MDEA), monoethanolamine (MEA), N-methylpiperazine (MP), piperazine, piperidine, 2-(2-tert-butylaminoethoxy)ethanol (TBEE), triethanolamine (TEA), triisopropanolamine (TIA), tris, 2-(2-aminoethoxy)ethanol, 2-(2-tert-butylaminopropoxy)ethanol, 2-(2-tert-amylaminoethoxy)ethanol, 2-(2-isopropylaminopropoxy)ethanol, 2-(2-(l -methyl- l-ethylpropylamino)ethoxy)ethanol, and mixtures thereof. In some embodiments, the amine compound is selected from: AMP, MEA, MDEA, TIA, and mixtures thereof. Further, in some embodiments the suitable conditions comprise an amine compound at a concentration of from about 1 M to about 10 M, from about 2 M to about 8 M, from about 2.5 M to about 6.5 M, from about 3 M to about 5 M, at least about 2 M, at least about 3 M, at least about 4.2 M, or at least about 5 M.

[0084] Solutions of amine compounds used for carbon dioxide absorption from gas streams typically are used at elevated temperatures. Accordingly, in some embodiments the improved property is stability in the presence of an amine compound and the suitable conditions comprise a solution temperature of from about 40°C to about 110°C, from about 40°C to about 90°C, from about 40°C to about 80°, from about 40°C to about 70°C, or from about 40°C to about 60°C.

[0085] Solutions containing ammonia that are used for carbon dioxide absorption from gas streams can be used at either or both chilled temperatures *(e.g.,* for absorption) and elevated temperatures *(e.g.,* for desorption of carbon dioxide). Accordingly, in some embodiments, the improved property is stability in ammonia and the suitable conditions comprise a solution temperature of from about 0°C to about 20°C, from about 0°C to about 10°C, from about 5°C to about 15°C, from about 8°C to about

12°C, less than about 15°C, or less than about 10°C. Further, in some embodiments the suitable conditions comprise an ammonia concentration of about 1 M to about 8 M, from about 2 M to about 7 M, from about 3 M to about 6 M, at least about 1 M, at least about 2 M, at least about 3 M, at least about 4 M, or at least about 5 M, or at least about 5.6 M.

[0086] Some solutions for the absorption of carbon dioxide from gas streams include high concentrations of carbonate ion ($C0_3^{2-}$). Typically, the carbonate ion is provided in the form of potassium carbonate (K_2C0_3) or sodium carbonate (Na_2C0_3). Accordingly, in some embodiments of the recombinant carbonic anhydrase polypeptides, the improved property is increased stability in solution comprising carbonate ion under suitable conditions, wherein the suitable conditions comprise a solution comprising carbonate ion at a concentration of about 0.1 M CO_3^{2-} to about 5 M $C0_3^{2-}$, from about 0.2 M $C0_3^{2-}$ to about 4 M $C0_3^{2-}$, or from about 0.3 M $C0_3^{2-}$ to about 3 M $C0_3^{2-}$.

[0087] The present disclosure also contemplates a recombinant carbonic anhydrase polypeptide having at least 1.3-fold, at least 1.5-fold, at least 2-fold, at least 3-fold, or at least 5-fold increased stability relative to the polypeptide of SEQ ID NO: 2, wherein the recombinant polypeptide comprises an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 2, and further comprises a set of amino acid residue differences as compared to SEQ ID NO:2, wherein the amino acid differences are based on locations or regions in the structure of reference polypeptide *(e.g.,* SEQ ID NO: 2) and/or the associated functional properties. Accordingly, referring to Table 3, a recombinant carbonic anhydrase polypeptide of the present disclosure can include an amino acid substitution at a particular residue at a location in the structure of the reference polypeptide as identified in Table 3. Exemplary substitutions at relevant locations include those identified in Tables 2A and 2B.

Table 3

Position	Structural Location
X2	Surface Exposed
X3	Surface Exposed
X4	Surface Exposed
X6	Surface Exposed
X7	Surface Exposed
X8	Surface Exposed
X12	Surface Exposed
X13	Buried
X14	Buried
X15	Surface Exposed
X16	Surface Exposed
X17	Buried
X18	Surface Exposed
X19	Surface Exposed
X20	Buried
X21	Buried
X22	Surface Exposed

X23	Surface Exposed
X24	Buried
X25	Surface Exposed
X26	Surface Exposed
X27	Surface Exposed
X28	Surface Exposed
X30	Surface Exposed
X36	Surface Exposed
X37	Surface Exposed
X38	Surface Exposed
X39	Buried
X41	Surface Exposed
X42	Buried
X43	Surface Exposed
X44	Surface Exposed
X46	Active Site - Outer Sphere
X47	Surface Exposed
X48	Buried
X50	Buried
X51	Buried
X52	Buried
X53	Buried
X54	Active Site - Outer Sphere - Buried
X55	Metal Coordinating - Buried
X56	Active Site - Outer Sphere - Buried
X57	Metal Coordinating - Buried
X58	Active Site - Outer Sphere - Buried
X59	Active Site - Outer Sphere - Buried
X60	Active Site - Outer Sphere - Buried
X61	Buried
X62	Buried
X63	Buried
X64	Buried
X65	Buried
X66	Buried
X67	Buried
X68	Buried
X69	Buried
X70	Buried
X71	Buried
X72	Buried
X73	Buried
X74	Active Site - Outer Sphere - Buried
X75	Buried
X76	Active Site - Outer Sphere -Buried
X77	Active Site - Outer Sphere - Buried
X78	Active Site - Outer Sphere - Buried
X79	Active Site - Inner Sphere - Buried
X80	Active Site - Inner Sphere - Buried
X81	Active Site - Outer Sphere - Buried
X82	Active Site - Outer Sphere - Buried
X83	Buried
X84	Buried

X86	Buried
X87	Buried
X88	Buried
X89	Buried
X90	Active Site - Outer Sphere - Buried
X91	Buried
X92	Buried
X93	Active Site - Outer Sphere
X94	Active Site - Outer Sphere - Buried
X95	Buried
X97	Active Site - Outer Sphere - Surface Exposed
X98	Active Site - Outer Sphere - Surface Exposed
X100	Buried
X101	Buried
X102	Buried
X103	Buried
X104	Buried
X105	Buried
X106	Active Site - Outer Sphere - Buried
X107	Active Site - Outer Sphere - Buried
X108	Metal Coordinating - Buried
X109	Active Site - Outer Sphere
X 110	Active Site - Outer Sphere - Surface Exposed
X I 11	- Metal Coordinating - Buried
X I 12	Active Site - Inner Sphere
X 113	Active Site - Inner Sphere - Buried
X 114	Active Site - Outer Sphere - Buried
X 115	Active Site - Outer Sphere - Surface Exposed
X 116	Active Site - Outer Sphere
X 117	Active Site - Outer Sphere - Buried
X 119	Surface Exposed
X120	Buried
X122	Surface Exposed
X123	Surface Exposed
X125	Buried
X126	Surface Exposed
X129	Surface Exposed
X130	Buried
X131	Buried
X132	Buried
X133	Buried
X135	Active Site - Outer Sphere - Buried
X136	Buried
X137	Buried
X139	Buried
X142	Buried
X143	Buried
X144	Buried
X146	Buried
X147	Surface Exposed
X148	Surface Exposed
X149	Surface Exposed
X150	Surface Exposed

X151	Buried
X152	Surface Exposed
X153	Surface Exposed
X154	Surface Exposed
X155	Surface Exposed
X156	Surface Exposed
X157	Buried
X158	Active Site - Outer Sphere - Buried
X160	Buried
X161	Active Site - Outer Sphere - Buried
X162	Buried
X163	Surface Exposed
X165	Buried
X166	Buried
X169	Buried
X172	Buried
X173	Buried
X176	Buried
X177	Surface Exposed
X178	Surface Exposed
X181	Surface Exposed
X182	Surface Exposed
X184	Surface Exposed
X185	Buried
X186	Buried
X187	Buried
X188	Buried
X189	Buried
X190	Buried
X191	Buried
X192	Active Site - Outer Sphere - Buried
X193	Surface Exposed
X194	Active Site - Outer Sphere - Buried
X195	Surface Exposed
X196	Surface Exposed
X197	Buried
X198	Surface Exposed
X199	Buried
X200	Surface Exposed
X201	Surface Exposed
X202	Surface Exposed
X203	Buried
X204	Surface Exposed
X205	Surface Exposed
X207	Surface Exposed
X208	Surface Exposed
X209	Surface Exposed

"inner sphere" - residue has an atom within 4.5 angstroms of the bound metal at active site.

"outer sphere" - residue within 4.5 angstroms of an inner sphere residue.

[0088] In some embodiments, any of the recombinant carbonic anhydrase polypeptides having at least 1.3-fold, at least 1.5-fold, at least 2-fold, at least 3-fold, or at least 5-fold increased stability relative to the polypeptide of SEQ ID NO: 2 and an amino acid sequence having at least 80%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity to SEQ ID NO: 2, can further comprise at least one amino acid residue difference selected from each of at least two of the following seven sets *(i.e.,* (a) through (g)) of amino acid residue differences: (a) X2R; X2T; X4F; (b) X121H; X121K; X121L; X121Q; X121T; X121V; X121W; X144A; X144L; X178G; (c) X139H; X139K; X139M; (d) X30A; X30L; X30Q; X30R; X40L; X40W; X68A; X96A; X96C; X96E; X96K; X I 19K; X I 19L; X I 19M; X I 19T; X120R; (e) X35R; X124G; X147E; X147F; X147G; X147T; X159H; X159R; (f) X31P; and (g) X56S; X84N; X84Q; X84S. In some embodiments, the recombinant carbonic anhydrase polypeptide amino acid sequence comprises one amino acid residue difference selected from each of at least two, three, four, five, six, or all seven of the sets of amino acid residue differences.

[0089] Although naturally occurring β-class carbonic anhydrases have been found that exhibit relatively high thermostability *(e.g.,* β-class carbonic anhydrase from *Methanobacterium thermoautotrophicum),* the β-class enzymes characterized to date have exhibited significantly lower specific activity in catalyzing the hydration of CO_2 *(e.g.,* reaction of Scheme 1) than the known alpha-class carbonic anhydrases *(e.g.,* alpha-class human CAII). For example, in a bicarbonate dehydration assay at pH 8.0, 25°C, the β-class CA from *M. thermoautotrophicum* exhibits less than 4% of the specific activity of the alpha-class human CAII. The actual difference is likely even greater because the initial rate of the human CAII was too fast to measure in this particular assay.

[0090] It is a surprising discovery of the present disclosure that the β-class carbonic anhydrase from *Desulfovibrio vulgaris* exhibits a high specific activity comparable to alpha-class enzymes and also exhibits high thermostability. For example, in the same bicarbonate dehydration assay comparison to the alpha-class human CAII, the specific activity of the β-class carbonic anhydrase from *D. vulgaris* of SEQ ID NO: 2 was found to be 84% - more than 20-fold greater than the β-class CA from *M. thermoautotrophicum.*

[0091] An analysis of the amino acid sequences of other naturally occurring β-class homologs having more than 40% identity to SEQ ID NO: 2 show that approximately 85% have a valine at position X60. The β-class carbonic anhydrase from *D. vulgaris* of SEQ ID NO: 2 has alanine at position X60. Structurally, the alanine at position X60 of SEQ ID NO: 2 resides just outside the metal binding site but contacts three of the four zinc coordinating residues C55, D57, and H108. Without being bound by mechanism, the structure-function correlation between the alanine at position X60 so close to the metal binding site and increased beta-class specific activity suggests that the volumetric change resulting from alanine rather than valine at position X60 results in greater active site flexibility, which

in turn results in the greater catalytic efficiency of the β-class carbonic anhydrase from *D. vulgaris* even at lower temperatures *(e.g.,* 5°C to 15°C).

[0092] In some embodiments, the present disclosure provides a beta-class carbonic anhydrase polypeptide capable of hydrating carbon dioxide in a solution comprising an amine compound or ammonia, wherein the polypeptide comprises an amino acid sequence at least about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to any one of beta carbonic anhydrases polypeptides of SEQ ID NO:2, 270, 272, 274, 276, or 278, and has an alanine residue at position **X60** relative to SEQ ID NO: 2. In some embodiments of the beta-class carbonic anhydrase polypeptides the amino acid sequence can further comprise one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2R; X2T; X4F; X4M; X15R; X16S; X30A; X30K; X30L; X30Q; X30R; X31P; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40W; X42A; X43M; X43V; X47R; X56S; X60C; X68A; X68G; X70I; X84N; X84Q; X84R; X84S; X86A; X95V; X96A; X96C; X96E; X96K; X97F; X I 19K; X I 19L; X I 19M; X I 19T; X120R; X121H; X121K; X121L; X121Q; X121T; X121V; X121W; X124G; X124R; X131L; X138L; X139H; X139K; X139M; X139Q; X142L; X144A; X144L; X145C; X145F; X145W; X147E; X147F; X147G; X147T; X148A; X148T; X159H; X159R; X159V; X168E; X170F; X178G; X200R; X213E; X213Q; X219T; X221C; X222C; and X223C.

[0093] In addition to the residue positions specified above, any of the recombinant carbonic anhydrase polypeptides disclosed herein can further comprise other residue differences relative to SEQ ID NO:2 at other residue positions. Residue differences at these other residue positions provide for additional variations in the amino acid sequence without adversely affecting the ability of the recombinant carbonic anhydrase polypeptide to carry out the hydration of carbon dioxide to bicarbonate and/or increased stability relative to the polypeptide of SEQ ID NO: 2. In some embodiments, the polypeptides can have additionally 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-1 1, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-26, 1-30, 1-35, 1-40 residue differences at other amino acid residue positions as compared to the reference sequence. In some embodiments, the number of differences can be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 1 1, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35, and 40 residue differences at other residue positions. The residue difference at these other positions can include conservative changes or non-conservative changes. In some embodiments, the residue differences can comprise conservative substitutions and non-conservative substitutions as compared to the wild-type carbonic anhydrase of SEQ ID NO: 2.

[0094] In some embodiments, the present disclosure provides recombinant carbonic anhydrase polypeptides that comprise deletions of the recombinant carbonic anhydrase polypeptides expressly described herein. Thus, for each and every embodiment comprising an amino acid sequence, there is another embodiment comprising a sequence having one or more amino acid deletions, 2 or more

amino acid deletions, 3 or more amino acid deletions, 4 or more amino acid deletions, 5 or more amino acid deletions, 6 or more amino acid deletions, 8 or more amino acid deletions, 10 or more amino acid deletions, 15 or more amino acid deletions, or 20 or more amino acid deletions, up to 10% of the total number of amino acids deleted, up to 20% of the total number of amino acids deleted, as long as the functional activity of the polypeptide with respect to the hydration of carbon dioxide to bicarbonate with increased stability is present. In some embodiments, the deletions can comprise, 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-1 1, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-26, 1-30, 1-35, or 1-40 amino acid residues. In some embodiments, the number of deletions can be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35, or 40 amino acids. In some embodiments, the deletions can comprise deletions of 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, or 20 amino acid residues.

[0095] In some embodiments, the polypeptides can comprise fragments of the recombinant carbonic anhydrase polypeptides described herein. In some embodiments, the fragments can have about 80%, 90%, 95%, 98%, and 99% of the full-length polypeptide, *e.g.,* the polypeptide of SEQ ID NO:2, as long as the functional activity of the polypeptide with respect to the hydration of carbon dioxide to bicarbonate with increased stability is present.

[0096] In some embodiments, the polypeptides of the disclosure can be in the form of fusion polypeptides in which the recombinant carbonic anhydrase polypeptides are fused to other polypeptides, such as, by way of example and not limitation, antibody tags *(e.g.,* myc epitope), purifications sequences *(e.g.,* His tags for binding to metals), and cell localization signals *(e.g.,* secretion signals). Thus, the recombinant carbonic anhydrase polypeptides described herein can be used with or without fusions to other polypeptides.

[0097] The polypeptides described herein are not restricted to the naturally-occurring genetically encoded L-amino acids but also include the D-enantiomers of the genetically-encoded amino acids. In addition to the genetically encoded amino acids, the polypeptides described herein may be comprised, either in whole or in part, of naturally-occurring and/or synthetic non-encoded amino acids that are known in the art (see, *e.g.,* the various amino acids provided in Fasman, 1989, CRC Practical Handbook of Biochemistry and Molecular Biology, CRC Press, Boca Raton, FL, at pp. 3-70 and the references cited therein, all of which are incorporated by reference). For example, conformationally constrained non-encoded amino acids of which the polypeptides described herein may be composed include: N-methyl amino acids (L-configuration); l-aminocyclopent-(2 or 3)-ene-4-carboxylic acid; pipecolic acid; azetidine-3-carboxylic acid; homoproline (hPro); and l-aminocyclopentane-3-carboxylic acid. Additionally, those of skill in the art will recognize that amino acids bearing side chain protecting groups may also comprise the polypeptides described herein - *e.g.,* Arg(tos), Cys(methylbenzyl), Cys (nitropyridinesulfenyl), Glu(8-benzylester), Gln(xanthyl), Asn(N-8-xanthyl),

His(bom), His(benzyl), His(tos), Lys(fmoc), Lys(tos), Ser(O-benzyl), Thr (O-benzyl) and Tyr(0-benzyl).

[0098] As described above the various modifications introduced into the naturally occurring polypeptide to generate an engineered carbonic anhydrase enzyme can be targeted to a specific property of the enzyme.

5.3. Polynucleotides Encoding Engineered Carbonic Anhydrases

[0099] In another aspect, the present disclosure provides polynucleotides encoding the engineered carbonic anhydrase enzymes. The polynucleotides may be operatively linked to one or more heterologous regulatory sequences that control gene expression to create a recombinant polynucleotide capable of expressing the polypeptide. Expression constructs containing a heterologous polynucleotide encoding the engineered carbonic anhydrase can be introduced into appropriate host cells to express the corresponding carbonic anhydrase polypeptide.

[0100] Because of the knowledge of the codons corresponding to the various amino acids, availability of a protein sequence provides a description of all the polynucleotides capable of encoding the subject. The degeneracy of the genetic code, where the same amino acids are encoded by alternative or synonymous codons allows an extremely large number of nucleic acids to be made, all of which encode the improved carbonic anhydrase enzymes disclosed herein. Thus, having identified a particular amino acid sequence, those skilled in the art could make any number of different nucleic acids by simply modifying the sequence of one or more codons in a way which does not change the amino acid sequence of the protein. In this regard, the present disclosure specifically contemplates each and every possible variation of polynucleotides that could be made by selecting combinations based on the possible codon choices, and all such variations are to be considered specifically disclosed for any polypeptide disclosed herein, including the amino acid sequences presented in Tables 2A and 2B.

[0101] In some embodiments, the polynucleotide comprises a nucleotide sequence encoding a recombinant carbonic anhydrase polypeptide with an amino acid sequence that has at least about 80% or more sequence identity, at least 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identity, or more sequence identity to any of the engineered carbonic anhydrase polypeptides described herein, *i.e.,* a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 240, 242, 244, 246, 248,

250, 252, 254, 256, 258, 260, 262, 264, 266, and 268. Exemplary polynucleotides encoding the engineered carbonic anhydrase are selected from SEQ ID NO: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 181, 183, 185, 187, 189, 191, 193, 195, 197, 199, 201, 203, 205, 207, 209, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 239, 241, 243, 245, 247, 249, 251, 253, 255, 257, 259, 261, 263, 265, 267, 269, 271, 273, 275, 277, 279, 281, 283, and 285.

[0102] In some embodiments, the polynucleotides encoding the engineered carbonic anhydrases are capable of hybridizing under highly stringent conditions to a polynucleotide comprising SEQ ID NO: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 181, 183, 185, 187, 189, 191, 193, 195, 197, 199, 201, 203, 205, 207, 209, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 239, 241, 243, 245, 247, 249, 251, 253, 255, 257, 259, 261, 263, 265, 267, 269, 271, 273, 275, 277, 279, 281, 283, and 285. These polynucleotides encode some of the recombinant carbonic anhydrase polypeptides represented by the amino acid sequences listed in Tables 2A and 2B.

[0103] In various embodiments, the codons are preferably selected to fit the host cell in which the recombinant carbonic anhydrase polypeptide is being produced. For example, preferred codons used in bacteria are used to express the gene in bacteria; preferred codons used in yeast are used for expression in yeast; and preferred codons used in mammals are used for expression in mammalian cells. For example, the polynucleotide of SEQ ID NO: 1 could be codon optimized for expression in *E. coli,* but otherwise encode the naturally occurring carbonic anhydrase of *Desulfovibrio vulgaris.*

[0104] In some embodiments, all codons need not be replaced to optimize the codon usage of the recombinant carbonic anhydrase polypeptide since the natural sequence will comprise preferred codons and because use of preferred codons may not be required for all amino acid residues. Consequently, codon optimized polynucleotides encoding the carbonic anhydrase enzymes may contain preferred codons at about 40%, 50%, 60%, 70%, 80%, or greater than 90% of codon positions of the full length coding region.

[0105] In other embodiments, the polynucleotides comprise polynucleotides that encode the recombinant carbonic anhydrase polypeptide described herein but have about 80% or more sequence identity, about 85% or more sequence identity, about 90% or more sequence identity, about 95% or

more sequence identity, about 98% or more sequence identity, or 99% or more sequence identity at the nucleotide level to a reference polynucleotide encoding an engineered carbonic anhydrase.

[0106] An isolated polynucleotide encoding an improved carbonic anhydrase polypeptide may be manipulated in a variety of ways to provide for expression of the polypeptide. Manipulation of the isolated polynucleotide prior to its insertion into a vector may be desirable or necessary depending on the expression vector. The techniques for modifying polynucleotides and nucleic acid sequences utilizing recombinant DNA methods are well known in the art. Guidance is provided in Sambrook *et al.*, 2001, Molecular Cloning: A Laboratory Manual, 3rd Ed., Cold Spring Harbor Laboratory Press; and Current Protocols in Molecular Biology, Ausubel. F. ed., Greene Pub. Associates, 1998, updates to 2006.

[0107] For bacterial host cells, suitable promoters for directing transcription of the nucleic acid constructs of the present disclosure, include the promoters obtained from the *E. coli lac* operon, *Streptomyces coelicolor* agarase gene (dagA), *Bacillus subtilis* levansucrase gene (sacB), *Bacillus licheniformis* alpha-amylase gene (amyL), *Bacillus stearothermophilus* maltogenic amylase gene (amyM), *Bacillus amyloliquefaciens* alpha-amylase gene (amyQ), *Bacillus licheniformis* penicillinase gene (penP), *Bacillus subtilis* xylA and xylB genes, and prokaryotic β-lactamase gene (Villa-Kamaroff *et al.*, 1978, Proc. Natl Acad. Sci. USA 75: 3727-3731), as well as the tac promoter (DeBoer *et al.*, 1983, Proc. Natl Acad. Sci. USA 80: 21-25). Further promoters are known in the art (see *e.g.*, "Useful proteins from recombinant bacteria" in Scientific American, 1980, 242:74-94; and in Sambrook *et al., supra.)*

[0108] For filamentous fungal host cells, suitable promoters for directing the transcription of the nucleic acid constructs of the present disclosure include promoters obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Rhizomucor miehei* aspartic proteinase, *Aspergillus niger* neutral alpha-amylase, *Aspergillus niger* acid stable alpha-amylase, *Aspergillus niger or Aspergillus awamori* glucoamylase (glaA), *Rhizomucor miehei* lipase, *Aspergillus oryzae* alkaline protease, *Aspergillus oryzae* triose phosphate isomerase, *Aspergillus nidulans* acetamidase, and *Fusarium oxysporum* trypsin-like protease (WO 96/00787), as well as the NA2-tpi promoter (a hybrid of the promoters from the genes for *Aspergillus niger* neutral alpha-amylase and *Aspergillus oryzae* triose phosphate isomerase), and mutant, truncated, and hybrid promoters thereof.

[0109] In a yeast hosts, useful promoters can be from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* galactokinase (GAL1), *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH2/GAP), and *Saccharomyces cerevisiae* 3-phosphoglycerate kinase. Other useful promoters for yeast host cells are described by Romanos *et al*, 1992, Yeast 8:423-488.

[0110] The control sequence may also be a suitable transcription terminator sequence, a sequence recognized by a host cell to terminate transcription. The terminator sequence is operably linked to the 3' terminus of the nucleic acid sequence encoding the polypeptide. Any terminator which is functional in the host cell of choice may be used in the present invention. For example, exemplary transcription terminators for filamentous fungal host cells can be obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* glucoamylase, *Aspergillus nidulans* anthranilate synthase, *Aspergillus niger* alpha-glucosidase, and *Fusarium oxysporum* trypsin-like protease. Exemplary terminators for yeast host cells can be obtained from the genes for *Saccharomyces cerevisiae* enolase, *Saccharomyces cerevisiae* cytochrome C (CYC1), and *Saccharomyces cerevisiae* glyceraldehyde-3 -phosphate dehydrogenase. Other useful terminators for yeast host cells are described by Romanos *et al.,* 1992, *supra.*

[0111] The control sequence may also be a suitable leader sequence operably linked to the 5' terminus of the nucleic acid sequence encoding the polypeptide. Any leader sequence that is functional in the host cell of choice may be used. Exemplary leaders for filamentous fungal host cells are obtained from the genes for *Aspergillus oryzae* TAKA amylase and *Aspergillus nidulans* triose phosphate isomerase. Suitable leaders for yeast host cells are obtained from the genes for *Saccharomyces cerevisiae* enolase (ENO-1), *Saccharomyces cerevisiae* 3-phosphoglycerate kinase, *Saccharomyces cerevisiae* alpha-factor, and *Saccharomyces cerevisiae* alcohol dehydrogenase/glyceraldehyde-3-phosphate dehydrogenase (ADH2/GAP).

[0112] The control sequence may also be a polyadenylation sequence operably linked to the 3' terminus of the nucleic acid sequence which, when transcribed, is recognized by the host cell as a signal to add polyadenosine residues to transcribed mRNA. Any polyadenylation sequence which is functional in the host cell of choice may be used in the present invention. Exemplary polyadenylation sequences for filamentous fungal host cells can be from the genes for *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* glucoamylase, *Aspergillus nidulans* anthranilate synthase, *Fusarium oxysporum* trypsin-like protease, and *Aspergillus niger* alpha-glucosidase. Useful polyadenylation sequences for yeast host cells are described by Guo and Sherman, 1995, Mol Cell Bio 15:5983-5990.

[0113] The control sequence may also be a signal peptide coding region that codes for an amino acid sequence linked to the amino terminus of an engineered carbonic anhydrase polypeptide and directs the encoded polypeptide into the cell's secretory pathway. The 5'-end of the coding sequence of the nucleic acid sequence may inherently contain a signal peptide coding region naturally linked in translation reading frame with the segment of the coding region that encodes the secreted polypeptide. Alternatively, the 5'-end of the coding sequence may contain a signal peptide coding region that is foreign to the coding sequence. The foreign signal peptide coding region may be required where the coding sequence does not naturally contain a signal peptide coding region. In some embodiments, the foreign signal peptide coding region may simply replace the natural signal peptide coding region in

order to enhance secretion of the polypeptide. However, any signal peptide coding region which directs the expressed polypeptide into the secretory pathway of a host cell of choice may be used in the present invention. Accordingly, an engineered carbonic anhydrase polypeptide of the invention can be operably linked to a signal sequence derived from a bacterial species such as a signal sequence derived from a *Bacillus (e.g., B. stearothermophilus , B. licheniformis , B. subtilis,* and *B. megaterium).* Exemplary signal peptide coding regions for bacterial host cells are the signal peptide coding regions obtained from the genes for *Bacillus* NC1B 11837 maltogenic amylase, *Bacillus stearothermophilus* alpha-amylase, *Bacillus licheniformis* subtilisin, *Bacillus licheniformis* beta-lactamase, *Bacillus stearothermophilus* neutral proteases (nprT, nprS, nprM), *Bacillus megaterium* enzymes (nprM, yngK, penG), and *Bacillus subtilis* prsA. Further signal peptides are described by Simonen and Palva, 1993, Microbiol Rev 57: 109-137. Exemplary signal peptide coding regions for filamentous fungal host cells can be the signal peptide coding regions obtained from the genes for *Aspergillus oryzae* TAKA amylase, *Aspergillus niger* neutral amylase, *Aspergillus niger glucoamylase, Rhizomucor miehei* aspartic proteinase, *Humicola insolens* cellulase, and *Humicola lanuginosa* lipase. Exemplary signal peptides for yeast host cells can be from the genes for *Saccharomyces cerevisiae* alpha-factor and *Saccharomyces cerevisiae* invertase. Other useful signal peptide coding regions are described by Romanos *et al.,* 1992, *supra.*

[0114] The control sequence may also be a propeptide coding region that codes for an amino acid sequence positioned at the amino terminus of a polypeptide. The resultant polypeptide is known as a pro-enzyme or pro-polypeptide (or a zymogen in some cases). A pro-polypeptide is generally inactive and can be converted to a mature active polypeptide by catalytic or autocatalytic cleavage of the propeptide from the pro-polypeptide. The pro-peptide coding region may be obtained from the genes for *Bacillus subtilis* alkaline protease (aprE), *Bacillus subtilis* neutral protease (nprT), *Saccharomyces cerevisiae* alpha-factor, *Rhizomucor miehei* aspartic proteinase, and *Myceliophthora thermophila* lactase (WO 95/33836).

[0115] Where both signal peptide and propeptide regions are present at the amino terminus of a polypeptide, the propeptide region is positioned next to the amino terminus of a polypeptide and the signal peptide region is positioned next to the amino terminus of the propeptide region.

[0116] It may also be desirable to add regulatory sequences, which allow the regulation of the expression of the polypeptide relative to the growth of the host cell. Examples of regulatory systems are those which cause the expression of the gene to be turned on or off in response to a chemical or physical stimulus, including the presence of a regulatory compound. In prokaryotic host cells, suitable regulatory sequences include the *lac, tac,* and *trp* operator systems. In yeast host cells, suitable regulatory systems include, as examples, the ADH2 system or GAL1 system. In filamentous fungi, suitable regulatory sequences include the TAKA alpha-amylase promoter, *Aspergillus niger* glucoamylase promoter, and *Aspergillus oryzae* glucoamylase promoter.

[0117] Other examples of regulatory sequences are those which allow for gene amplification. In eukaryotic systems, these include the dihydrofolate reductase gene, which is amplified in the presence of methotrexate, and the metallothionein genes, which are amplified with heavy metals. In these cases, the nucleic acid sequence encoding the carbonic anhydrase polypeptide of the present invention would be operably linked with the regulatory sequence.

[0118] In another embodiment, the present disclosure is also directed to a recombinant expression vector comprising a polynucleotide encoding an engineered carbonic anhydrase polypeptide or a variant thereof, and one or more expression regulating regions such as a promoter and a terminator, a replication origin, etc., depending on the type of hosts into which they are to be introduced. The various nucleic acid and control sequences described above may be joined together to produce a recombinant expression vector which may include one or more convenient restriction sites to allow for insertion or substitution of the nucleic acid sequence encoding the polypeptide at such sites. Alternatively, the nucleic acid sequence of the present disclosure may be expressed by inserting the nucleic acid sequence or a nucleic acid construct comprising the sequence into an appropriate vector for expression. In creating the expression vector, the coding sequence is located in the vector so that the coding sequence is operably linked with the appropriate control sequences for expression.

[0119] The recombinant expression vector may be any vector *(e.g.,* a plasmid or virus), which can be conveniently subjected to recombinant DNA procedures and can bring about the expression of the polynucleotide sequence. The choice of the vector will typically depend on the compatibility of the vector with the host cell into which the vector is to be introduced. The vectors may be linear or closed circular plasmids. The expression vector may be an autonomously replicating vector, *i.e.,* a vector that exists as an extrachromosomal entity, the replication of which is independent of chromosomal replication, *e.g.,* a plasmid, an extrachromosomal element, a minichromosome, or an artificial chromosome. The vector may contain any means for assuring self-replication. Alternatively, the vector may be one which when introduced into the host cell, is integrated into the genome and replicated together with the chromosome(s) into which it has been integrated. Furthermore, a single vector or plasmid or two or more vectors or plasmids which together contain the total DNA to be introduced into the genome of the host cell, or a transposon may be used.

[0120] An expression vector of the present disclosure preferably contains one or more selectable markers, which permit easy selection of transformed cells. Examples of bacterial selectable markers are the dal genes from *Bacillus subtilis* or *Bacillus licheniformis,* or markers, which confer antibiotic resistance such as ampicillin, kanamycin, chloramphenicol or tetracycline resistance. Suitable markers for yeast host cells are ADE2, HIS3, LEU2, LYS2, MET3, TRP1, and URA3. Selectable markers for use in a filamentous fungal host cell include, but are not limited to, amdS (acetamidase), argB (ornithine carbamoyltransferase), bar (phosphinothricin acetyltransferase), hph (hygromycin phosphotransferase), niaD (nitrate reductase), pyrG (orotidine-5'-phosphate decarboxylase), sC

(sulfate adenyltransferase), and trpC (anthranilate synthase), as well as equivalents thereof. Embodiments for use in an *Aspergillus* cell include the amdS and pyrG genes of *Aspergillus nidulans* or *Aspergillus oryzae* and the bar gene of *Streptomyces hygroscopicus*.

[0121] The expression vectors of the present disclosure also can contain an element(s) that permits integration of the vector into the host cell's genome or autonomous replication of the vector in the cell independent of the genome. For integration into the host cell genome, the vector may rely on the nucleic acid sequence encoding the polypeptide or any other element of the vector for integration of the vector into the genome by homologous or non-homologous recombination. Alternatively, the expression vector may contain additional nucleic acid sequences for directing integration by homologous recombination into the genome of the host cell. The additional nucleic acid sequences enable the vector to be integrated into the host cell genome at a precise location(s) in the chromosome(s). The integrational elements may be any sequence that is homologous with the target sequence in the genome of the host cell. Furthermore, the integrational elements may be non-encoding or encoding nucleic acid sequences.

[0122] In some embodiments, the vector may further comprise an origin of replication enabling the vector to replicate autonomously in the host cell in question. Examples of bacterial origins of replication are P15A ori or the origins of replication of plasmids pBR322, pUC19, pACYC177 (which plasmid has the P15A ori), or pACYC184 permitting replication in *E. coli*, and pUBl 10, pE194, pTA1060, or ρAMβI permitting replication in *Bacillus*. Examples of origins of replication for use in a yeast host cell are the 2 micron origin of replication, ARS 1, ARS4, the combination of ARS 1 and CEN3, and the combination of ARS4 and CEN6. The origin of replication may be one having a mutation which makes its function temperature-sensitive in the host cell (see, *e.g.*, Ehrlich, 1978, Proc Natl Acad Sci. USA 75:1433).

[0123] Many expression vectors useful with polynucleotides of the present disclosure are commercially available. Suitable commercial expression vectors include p3xFLAGTMTM expression vectors from Sigma-Aldrich Chemicals, St. Louis MO., which includes a CMV promoter and hGH polyadenylation site for expression in mammalian host cells and a pBR322 origin of replication and ampicillin resistance markers for amplification in *E. coli*. Other suitable expression vectors are *Bacillus megaterium* shuttle vector pMM1525 (Boca Scientific Inc. Boca Raton, FL), pBluescriptll SK(-) and pBK-CMV, which are commercially available from Stratagene, La Jolla CA, and plasmids which are derived from pBR322 (Gibco BRL), pUC (Gibco BRL), pREP4, pCEP4 (Invitrogen) or pPoly (Lathe *et al*, 1987, Gene 57:193-201).

5.4. Host Cells for Expression of Carbonic Anhydrase Polypeptides

[0124] In another aspect, the present disclosure provides a host cell comprising a polynucleotide encoding an improved carbonic anhydrase polypeptide of the present disclosure, the polynucleotide

being operatively linked to one or more control sequences for expression of the carbonic anhydrase enzyme in the host cell. Host cells for use in expressing the carbonic anhydrase polypeptides encoded by the expression vectors of the present invention are well known in the art and include but are not limited to, bacterial cells, such as *E. coli, Bacillus, Lactobacillus, Streptomyces* and *Salmonella typhimurium* cells; fungal cells, such as yeast cells *(e.g., Saccharomyces cerevisiae* or *Pichia pastoris* (ATCC Accession No. 201 178)); insect cells such as *Drosophila* S2 and *Spodoptera* Sf9 cells; animal cells such as CHO, COS, BHK, 293, and Bowes melanoma cells; and plant cells.

[0125] In some embodiments of the invention the host cell is a bacterial host cell of the *Bacillus* species, *e.g., B. thuringiensis, B. anthracis, B. megaterium, B. subtilis, B. lentus, B. circulans, B. pumilus, B. lautus, B.coagulans, B. brevis, B.firmus, B. alkaophius, B. licheniformis, B. clausii, B. stearothermophilus, B. halodurans* and *B. amyloliquefaciens*. Appropriate culture mediums and growth conditions for the above-described host cells are well known in the art.

[0126] Polynucleotides for expression of the carbonic anhydrase may be introduced into cells by various methods known in the art. Techniques include among others, electroporation, biolistic particle bombardment, liposome mediated transfection, calcium chloride transfection, and protoplast fusion. Various methods for introducing polynucleotides into cells will be apparent to the skilled artisan. In some embodiments, more than one copy of a polynucleotide sequence is inserted into a host cell to increase production of the gene product. An increase in the copy number of the nucleic acid sequence can be obtained by integrating at least one additional copy of the sequence into the host cell genome or by including an amplifiable selectable marker gene with the nucleic acid sequence where cells containing amplified copies of the selectable marker gene, and thereby additional copies of the nucleic acid sequence, can be selected for by cultivating the cells in the presence of the appropriate selectable agent.

[0127] An exemplary host cell is *Escherichia coli* W3110. An expression vector encoding an improved carbonic anhydrase of the present disclosure can be created by operatively linking a polynucleotide into the plasmid pCKl 10900 *(see,* U.S. application publication 20040137585) operatively linked to the lac promoter under control of the lacl repressor. The expression vector also contained the P15a origin of replication and the chloramphenicol resistance gene. Cells containing the subject polynucleotide in *Escherichia coli* W3110 were isolated by subjecting the cells to chloramphenicol selection. Another exemplary host cell is *Escherichia coli* BL21.

[0128] The disclosure also provides methods for producing the recombinant carbonic anhydrase polypeptides using a host cell. In some embodiments, the method for producing a recombinant carbonic anhydrase polypeptide comprises the steps of: (a) transforming a host cell with an expression vector polynucleotide encoding the recombinant carbonic anhydrase polypeptide; (b) culturing said transformed host cell under conditions whereby said recombinant carbonic anhydrase polypeptide is

produced by said host cell; and (c) recovering said recombinant carbonic anhydrase polypeptide from said host cells. In some embodiments, the methods of producing the recombinant carbonic anhydrase may be carried out wherein said expression vector comprises a secretion signal, and said cell is cultured under conditions whereby the recombinant carbonic anhydrase polypeptide is secreted from the cell. In some embodiments of the method, the expression vector comprises a polynucleotide encoding a secretion signal.

[0129] Recovery, isolation and purification of the recombinant carbonic anhydrase polypeptide may be carried out using standard methods known by the ordinary artisan such those as described further below.

5.5. Methods of Generating Engineered Carbonic Anhydrase Polypeptides.

[0130] In some embodiments, to make the improved carbonic anhydrase polynucleotides and polypeptides of the present disclosure, the naturally-occurring carbonic anhydrase enzyme that catalyzes the hydration reaction is obtained (or derived) from *Desulfovibrio vulgaris.* In some embodiments, the parent polynucleotide sequence is codon optimized to enhance expression of the carbonic anhydrase in a specified host cell. As an illustration, the parental polynucleotide sequence (SEQ ID NO: 1) encoding the wild-type carbonic anhydrase polypeptide of *Desulfovibrio vulgaris* (SEQ ID NO:2), can be assembled from oligonucleotides based upon that sequence or from oligonucleotides comprising a codon-optimized coding sequence for expression in a specified host cell, *e.g.,* an *E. coli* host cell. In one embodiment, the polynucleotide can be cloned into an expression vector, placing the expression of the carbonic anhydrase gene under the control of the lac promoter and lacl repressor gene. Clones expressing the active carbonic anhydrase in *E. coli* can be identified and the genes sequenced to confirm their identity.

[0131] The engineered carbonic anhydrase can be obtained by subjecting the polynucleotide encoding the naturally occurring carbonic anhydrase to mutagenesis and/or directed evolution methods, as discussed above. An exemplary directed evolution technique is mutagenesis and/or DNA shuffling as described in Stemmer, 1994, Proc Natl Acad Sci USA 91:10747-10751; WO 95/22625; WO 97/0078; WO 97/35966; WO 98/27230; WO 00/42651; WO 01/75767 and U.S. Pat. 6,537,746. Other directed evolution procedures that can be used include, among others, staggered extension process (StEP), in vitro recombination (Zhao *et al.,* 1998, Nat. Biotechnol. 16:258-261), mutagenic PCR (Caldwell *et al,* 1994, PCR Methods Appl. 3:S136-S140), and cassette mutagenesis (Black *et al,* 1996, Proc Natl Acad Sci USA 93:3525-3529).

[0132] Methodologies for screening and identifying polypeptides for desired activities are useful in the preparation of new compounds such as modified enzymes and/or new pharmaceuticals. Directed evolution can be used to discover or enhance activity of polypeptides of commercial interest. For example, if the activity of a known catalyst is insufficient for a commercial process, directed

evolution and/or other protein engineering technologies may be used to make appropriate improvements to the catalyst to improve activity on the substrate of interest. Improvements to process engineering can be developed to enhance an active enzyme and/or to optimize a microbe/enzyme for scaled-up production. Current methodologies are often limited by time and cost factors. In some instances, it may take months or years, at great expense, to find a new polypeptide with the desired activity, if one is ever found. Furthermore, the number of polypeptide variants that must be screened is often cumbersome. Thus, there is a long felt need for compositions and methods used to identify novel polypeptide variants having a desired activity.

[0133] Many methodologies directed to the design and/or identification of polypeptides having particular characteristics are known in the art. For example, methods for high-throughput screening arrays of clones in a sequential manner are presented in PCT Publication No. WO 01/32858; an in vitro selection method of screening a library of catalyst molecules is disclosed in PCT Publication No. WO 00/1 121 1; a screening method for identifying active peptides or proteins with improved performance is disclosed in PCT Publication No. WO 02/072876 and U.S. Patent Application Publication No. 2004/0132039; a methods for creating and screening transgenic organisms having desirable traits are disclosed in U.S. Patent No. 7,033,78 1; methods for making circularly permuted proteins and peptides having novel and/or enhanced functions with respect to a native protein or peptide are disclosed in PCT Publication No. WO 2006/086607; methods for preparing variants of a catalytic polypeptide are disclosed in U.S. Patent Application Publication No. 2003/0073 109; and methods for biopolymer engineering using a variant set to model sequence-activity relationships are disclosed in PCT Publication No. WO 2005/013090; each of which is incorporated herein by reference in its entirety.

[0134] The clones obtained following mutagenesis treatment are screened for engineered carbonic anhydrase having a desired improved enzyme property. Measuring enzyme activity from the expression libraries can be performed using the standard biochemistry technique of monitoring changes in pH, either directly or indirectly, as indicated in the Examples. Similarly, and as again demonstrated in the Examples, activity of the carbonic anhydrases of the disclosure may be measured using either the forward or reverse reactions depicted in Scheme 1. Where the improved enzyme property desired is thermal stability, enzyme activity may be measured after subjecting the enzyme preparations to a defined temperature for a defined period of time and measuring the amount of enzyme activity remaining after heat treatments. Clones containing a polynucleotide encoding a carbonic anhydrase are then isolated, sequenced to identify the nucleotide sequence changes (if any), and used to express the enzyme in a host cell.

[0135] Where the sequence of the engineered polypeptide is known, the polynucleotides encoding the enzyme can be prepared by standard solid-phase methods, according to known synthetic methods. In some embodiments, fragments of up to about 100 bases can be individually synthesized, then

joined *(e.g.,* by enzymatic or chemical litigation methods, or polymerase mediated methods) to form any desired continuous sequence. For example, polynucleotides and oligonucleotides of the invention can be prepared by chemical synthesis using, *e.g.,* the classical phosphoramidite method described by Beaucage *et al,* 1981, Tet Lett 22:1859-69, or the method described by Matthes *et al,* 1984, EMBO J. 3:801-05, *e.g.,* as it is typically practiced in automated synthetic methods. According to the phosphoramidite method, oligonucleotides are synthesized, *e.g.,* in an automatic DNA synthesizer, purified, annealed, ligated and cloned in appropriate vectors. In addition, essentially any nucleic acid can be obtained from any of a variety of commercial sources, such as The Midland Certified Reagent Company, Midland, TX, The Great American Gene Company, Ramona, CA, ExpressGen Inc. Chicago, IL, Operon Technologies Inc., Alameda, CA, and many others.

[0136] Engineered carbonic anhydrase enzymes expressed in a host cell can be recovered from the cells and or the culture medium using any one or more of the well known techniques for protein purification, including, among others, lysozyme treatment, sonication, filtration, salting-out, ultra-centrifugation, and chromatography. Suitable solutions for lysing and the high efficiency extraction of proteins from bacteria, such as *E. coli,* are commercially available under the trade name CelLytic BTM from Sigma-Aldrich of St. Louis MO.

[0137] Chromatographic techniques for isolation of the carbonic anhydrase polypeptide include, among others, reverse phase chromatography high performance liquid chromatography, ion exchange chromatography, gel electrophoresis, and affinity chromatography. Conditions for purifying a particular enzyme will depend, in part, on factors such as net charge, hydrophobicity, hydrophilicity, molecular weight, molecular shape, etc., and will be apparent to those having skill in the art.

[0138] In some embodiments, affinity techniques may be used to isolate the improved carbonic anhydrase enzymes. For affinity chromatography purification, any antibody which specifically binds the carbonic anhydrase polypeptide may be used. For the production of antibodies, various host animals, including but not limited to rabbits, mice, rats, *etc.,* may be immunized by injection with a polypeptide of the disclosure. The polypeptide may be attached to a suitable carrier, such as BSA, by means of a side chain functional group or linkers attached to a side chain functional group. Various adjuvants may be used to increase the immunological response, depending on the host species, including but not limited to Freund's (complete and incomplete), mineral gels such as aluminum hydroxide, surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, keyhole limpet hemocyanin, dinitrophenol, and potentially useful human adjuvants such as BCG (bacilli Calmette Guerin) and *Corynebacterium parvum.*

5.6. Methods of Using the β-Class Carbonic Anhydrase Polypeptides

[0139] The β-class carbonic anhydrase enzymes described herein can catalyze both the forward and reverse reactions depicted in Scheme 1. In certain embodiments, a carbonic anhydrase of the present disclosure can be used to hydrate carbon dioxide in the form of bicarbonate and a proton, which in turn, will be converted to carbonate and/or a mixture of bicarbonate and carbonate at an elevated pH. In other embodiments, a carbonic anhydrase of the disclosure can be used to dehydrate carbon dioxide by reaction at a relatively acidic pH, thereby catalyzing the release of hydrated CO_2 from solution.

[0140] Accordingly, in some embodiments the present disclosure provides methods for removing carbon dioxide from a gas stream *(e.g.,* capturing or extracting CO_2 gas) comprising the step of contacting the gas stream with a solution comprising a carbonic anhydrase polypeptide of the disclosure having an improved property *(e.g.,* increased activity, thermostability and/or solvent stability), whereby carbon dioxide from the gas stream is absorbed into the solution (e.g., $C0_2$ gas diffuses into solution and is hydrated to bicarbonate).

[0141] In other embodiments, the methods of the disclosure can comprise further steps of isolating and/or separately treating the solution comprising the absorbed carbon dioxide. In some embodiments, the carbon dioxide gas in the solution is desorbed *(i.e.,* stripped) by contacting the isolated solution with protons *(i.e.,* acidify) and a carbonic anhydrase polypeptide, which may be the same or different than the polypeptide used in the absorption step, thereby converting the hydrated carbon dioxide to carbon dioxide gas and water. In some embodiments, the desorption of carbon dioxide from this separate solution can be carried out at significantly higher temperatures, and/or under lower pressure *(e.g.,* vacuum) conditions that require a recombinant carbonic anhydrase having different stability characteristics. Thus, it is contemplated that the solution can be removed from contact with the gas stream *(e.g.,* isolated after some desired level of hydrated carbon dioxide is reached) and further treated with a carbonic anhydrase to convert the bicarbonate in solution into carbon dioxide gas, which is then released from the solution and sequestered, *e.g.,* into a pressurized chamber. In other embodiments, the further step of isolating the solution comprising the hydrated carbon dioxide is carried out and no further carbonic anhydrase polypeptide is added to the solution. Instead the solution which is enriched in bicarbonate *(i.e.,* hydrated carbon dioxide) can be used in processes that react with the bicarbonate to effectively sequester the carbon dioxide in another chemical form.

[0142] In some embodiments, the carbonic anhydrases and associated methods for removing *(e.g.,* extracting and sequestering) carbon dioxide from a gas stream disclosed herein can be used in existing systems that use a solution for absorbing carbon dioxide from *e.g.,* flue gas. Equipment, processes, and methods for carbon dioxide capture and sequestration using solutions into which carbon dioxide is absorbed *(i.e.,* captured by diffusing from gas stream into the liquid solution) and/or from which carbon dioxide is desorbed *(i.e.,* extracted by diffusing from liquid solution into gas phase) are

described in *e.g.,* U.S. Pat. Nos. 6,143,556, 6,524,843 B2, 7,176,017 B2, 7,596,952 B2, 7,641,717 B2, 7,579185 B2, 7,740,689 B2, 7,132,090 B2; U.S. Pat. Publ. Nos. 2007/0004023A1, 2007/0256559A1, 2009/0155889A1, 2010/0086983A1; PCT Publ. Nos. WO98/55210A1, WO2004/056455A1, WO2004/028667A1, WO2006/089423A1, WO2008/072979A1, WO2009/000025A1, WO2010/020017A1, WO2010/014773Al, WO2010/045689A1, each of which is hereby incorporated by reference herein.

[0143] Methods for linking (covalently or non-covalently) enzymes to solid-phase particles *{e.g.,* porous or non-porous beads, or solid supports) such that they retain activity for use in bioreactors are known in the art. Methods for treating a gas stream using immobilized enzymes are described in *e.g.,* U.S. patent no. 6,143,556, U.S. patent publication no. 2007/0004023A1, and PCT publications WO98/55210A1, WO2004/056455A1, and WO2004/028667A1, each of which is hereby incorporated by reference herein. Accordingly, in some embodiments, the methods for removing carbon dioxide from a gas stream can be carried out wherein the recombinant carbonic anhydrase polypeptide is immobilized on a surface, for example wherein the enzyme is linked to the surface of a solid-phase particle *(e.g.,* beads) in the solution. In some embodiments, the methods using immobilized polypeptides can be carried out wherein the method further comprises a step of isolating or separating the immobilized carbonic anhydrase from the solution. After separating the immobilized carbonic anhydrase from the solution, the solution can be treated to conditions that may inactivate the enzyme, e.g., desorption of CO_2 at high temperatures. Further, the separately retained immobilized enzyme can be added to another solution and reused.

[0144] In some embodiments, the methods of removing carbon dioxide from a gas stream can be carried out using a naturally occurring β-class carbonic anhydrase polypeptide of any one of SEQ ID NO: 2, 270, 272, 274, 276, and 278. In some embodiments, the methods can be carried out using the carbonic anhydrase polypeptide of SEQ ID NO: 2. In some embodiments, the methods can be carried out using a β-class carbonic anhydrase polypeptide, wherein the polypeptide comprises an amino acid sequence at least about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical SEQ ID NO:2 and has an alanine residue at position **X60** relative to SEQ ID NO: 2.

[0145] In some embodiments, the methods of removing carbon dioxide from a gas stream can be carried out wherein the carbonic anhydrase polypeptide comprises an amino acid sequence that is at least about 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to the amino acid sequence of SEQ ID NO:2, and comprises one or more amino acid differences relative to SEQ ID NO: 2 selected from the group consisting of: X2R; X2T; X4F; X4M; X15R; X16S; X22G; X30A; X30K; X30L; X30Q; X30R; X31P; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40Q; X40W; X42A; X43M; X43V; X47R; X56S; X60C; X60V; X68A; X68G; X70I; X84K; X84N; X84Q; X84R; X84S; X86A; X95V; X96A; X96C; X96E; X96K; X97F; X I 19K;

X I 19L; X I 19M; X I 19T; X120R; X121H; X121K; X121L; X121Q; X121T; X121V; X121W; X124G; X124R; X131L; X138F; X138L; X138W; X139H; X139K; X139M; X139Q; X142L; X143M; X144A; X144L; X145C; X145F; X145W; X147E; X147F; X147G; X147T; X148A; X148T; X157A; X159H; X159R; X159V; X168E; X170F; X178G; X200R; X207E; X207N; X213E; X213Q; X219T; X221C; X222C; X223C; and X223Q. The foregoing carbonic anhydrase polypeptides may further comprise additional modifications, including substitutions, deletions, insertions, or combinations thereof. The substitutions can be non-conservative substitutions, conservative substitutions, or a combination of non-conservative and conservative substitutions. In some embodiments, these carbonic anhydrase polypeptides can have optionally from about 1-2, 1-3, 1-4, 1-5, 1-6, 1-7, 1-8, 1-9, 1-10, 1-1 1, 1-12, 1-14, 1-15, 1-16, 1-18, 1-20, 1-22, 1-24, 1-25, 1-30, 1-35 or about 1-40 mutations at other amino acid residues. In some embodiments, the number of modifications can be 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 14, 15, 16, 18, 20, 22, 24, 26, 30, 35 or about 40 other amino acid residues.

[0146] In certain embodiments, the methods can be carried out using a recombinant carbonic anhydrase polypeptide of the present disclosure, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 240, 242, 244, 246, 248, 250, 252, 254, 256, 258, 260, 262, 264, 266, and 268.

[0147] In various embodiments, the methods of removing carbon dioxide from a gas stream using a carbonic anhydrase polypeptide disclosed herein may be carried out under a range of suitable conditions. Suitable conditions can be determined by routine experimentation that includes, but is not limited to, contacting the solution containing the carbonic anhydrase polypeptide with CO_2 at an experimental condition *(e.g.,* amine concentration, temperature, CO_2 loading) and then detecting the relevant activity *(e.g.,* rate of CO_2 absorption), for example, using the methods described in the Examples provided herein.

[0148] The ordinary artisan also will recognize that certain suitable conditions can be selected that favor the absorption of carbon dioxide into a solution *(e.g.,* via hydration of carbon dioxide to bicarbonate) and/or the desorption of carbon dioxide from a solution *(e.g.,* via dehydration of bicarbonate to carbon dioxide and water). The carbonic anhydrase polypeptides disclosed herein are biocatalysts having an improved property *(e.g.,* thermal stability, solvent stability, and/or base stability) that allows them to accelerate the absorption of carbon dioxide gas into a solution and/or accelerate subsequent desorption from the solution under a range of conditions.

[0149] In some embodiments, the method can be carried out wherein the carbonic anhydrase comprises the improved property at least 1.2-fold, at least 1.3-fold, at least 1.5-fold, at least 2-fold, at least 3-fold, at least 4-fold, at least 5-fold, at least 10-fold, or at least 25-fold increased activity of hydrating carbon dioxide or dehydrating bicarbonate under suitable conditions. Accordingly, in some embodiments, the suitable conditions used in the method can comprise a concentration of the carbonic anhydrase polypeptide of from about 0.1 g/L to about 10 g/L, about 0.25 g/L to about 7.5 g/L, about 0.5 g/L to about 5 g/L, less than 10 g/L, less than about 5 g/L, or less than about 2.5 g/L.

[0150] In some embodiments, the method is carried out wherein the carbonic anhydrase polypeptide is added to the solution in the form of a powder *(e.g.,* shake-flask powder, or DSP powder). The powder may contain the polypeptide in partially purified or a highly purified form. Alternatively, in embodiments, the methods can be carried out wherein the suitable conditions comprising adding whole cells transformed with a gene encoding the enzyme, and/or cell extracts and/or lysates of such cells. Whole cells transformed with a gene encoding the engineered carbonic anhydrase enzyme or cell extracts and/or lysates thereof, may be employed in a variety of different forms, including solid *(e.g.,* lyophilized, spray-dried, and the like) or semisolid *(e.g.,* a crude paste) forms. The cell extracts or cell lysates may be partially purified by precipitation (ammonium sulfate, polyethyleneimine, heat treatment or the like, followed by a desalting procedure prior to lyophilization *(e.g.,* ultrafiltration, dialysis, and the like). Any of the cell preparations may be stabilized by crosslinking using known crosslinking agents, such as, for example, glutaraldehyde or immobilization to a solid phase *(e.g.,* Eupergit C, and the like) or by the crosslinking of protein crystals or precipitated protein aggregate particles.

[0151] The ability of the carbonic anhydrase polypeptide to accelerate $C0_2$ absorption into or desorption from a solution can be affected by the mole ratio of $C0_2$ to other compounds already present in the solution, which is also referred to as the $C0_2$ loading of the solution and can be denoted by the mole ratio of $C0_2$ to the moles of the relevant $C0_2$ absorption mediating compound in the solution *(e.g.,* amine compound, ammonia), which is denoted by the term "a." The carbonic anhydrase polypeptides of the present disclosure can be used under a range of loading conditions which can be varied depending on the particular $C0_2$ absorption mediating compound used in the solution. Accordingly, the methods of the disclosure can be carried wherein the suitable conditions comprise a loading of solution $C0_2$ of from about a = 0 to about a = 0.7, from about a = 0.01 to about a = 0.6, from about a = 0.02 to about a = 0.5, from about a = 0.05 to about a = 0.4, from about a = 0.1 to about a = 0.4, from about a = 0.2 to about a = 0.3, less than about a = 0.7, less than about a = 0.5, or less than about a = 0.3.

[0152] In some embodiments where the method is carried out in the presence of an amine compound *(e.g.,* 4 M MDEA) the suitable conditions can comprise and a loading of solution $C0_2$ of from about a = 0 to about a = 0.6, from about a = 0.01 to about a = 0.5, from about a = 0.02 to about a = 0.4, from

about a = 0.05 to about a = 0.3, from about a = 0.1 to about a = 0.4, from about a = 0.2 to about a = 0.3, less than about a = 0.4, less than about a = 0.3, or less than about a = 0.2.

[0153] In some embodiments where the method is carried out in the presence of ammonia *(e.g.,* 10 wt% or 5.6 M NH_3) the suitable conditions can comprise a loading of solution $C0_2$ of from about a = 0 to about a = 0.7, from about a = 0.1 to about a = 0.7, from about a = 0.1 to about a = 0.5, from about a = 0.1 to about a = 0.3, from about a = 0.4 to about a = 0.7, from about a = 0.5 to about a = 0.7, less than about a = 0.7, less than about a = 0.5, or less than about a = 0.3.

[0154] Additionally, the CO_2 loading of the solution can change from "lean" to "rich" during the process as the CO_2 is absorbed, and then desorbed. Typically, the initial condition of the solution used in the method is "lean loading" *(e.g., a* = 0, or a = 0.01 to 0.02), and as the absorption proceeds the solution condition becomes "rich loading" *(e.g., a* = 0.2 to 0.5, or higher). As illustrated by the Examples herein, the acceleration of CO_2 absorption due to enzyme tends to be lower under "lean loading" conditions than under "rich loading" conditions. Further the loading conditions used for the method carried out in the presence of amine compounds tends to be lower than the loading used for the method carried out in the presence of ammonia. Accordingly, in some embodiments, the suitable conditions in the presence of an amine compound comprise a lean loading of solution CO_2 from about a = 0 to about a = 0.02 and a rich loading of solution CO_2 of from about a = 0.2 to about a = 0.5. However, in some embodiments, where the suitable conditions include the presence of ammonia, the loading can comprise a lean loading of solution CO_2 about a = 0.1 to about a = 0.3 and a rich loading of solution CO_2 of from about a = 0.5 to about a = 0.7.

[0155] Typically the gas streams from which CO_2 removal is desirable are at elevated temperatures, and upon contacting a solution, as in the method disclosed herein, heat is also transferred and the solution temperature also is elevated. This is particularly true in treating flue gas streams from coal-fired power plants. Accordingly, in some embodiments, the suitable conditions for carrying out the method comprise an elevated solution temperature. The presence of elevated temperature further underscores the importance of using thermostable carbonic anhydrase polypeptides such as those disclosed herein. Thus, in some embodiments the method is carried out wherein the suitable conditions comprise a solution temperature of from about 40°C to about 110°C, from about 40°C to about 90°C, from about 40°C to about 80°, from about 40°C to about 70°C, or from about 40°C to about 60°C.

[0156] The method of removing carbon dioxide disclosed herein involves contacting the gas stream with a solution comprising a carbonic anhydrase polypeptide. The present disclosure has illustrated the use of the method in solutions comprising a high concentration of an amine compound, ammonia, and carbonate ion. A range of other solutions comprising other compounds known to facilitate the

absorption of CO_2 from a gas stream, and it is contemplated that the present methods could be used with such solutions.

[0157] For capturing CO_2 from flue gas streams, solutions comprising a variety of different amine compounds are known. Such solutions comprising amine compounds that facilitate $C0_2$ absorption from a gas stream into a solution are described in *e.g.,* PCT Publ. No. WO2006/089423A1, U.S. Pat. No. 7,740,689 B2, or U.S. Pat. Publ. No. 2009/0155889A1, each which is hereby incorporated by reference herein. Accordingly, in some embodiments, the methods of removing carbon dioxide from a gas stream can be carried out wherein the solution comprises an amine compound, preferably an amine compound that exhibits improved thermodynamic and kinetic properties for the absorption of CO_2. Thus, in some embodiments of the methods, the suitable conditions comprise a solution comprising an amine compound, and the amine compound can be selected from: 2-(2-aminoethylamino)ethanol (AEE), 2-amino-2-hydroxymethyl-l,3-propanediol (AHPD), 2-amino-2-methyl- 1-propanol (AMP), diethanolamine (DEA), diisopropanolamine (DIPA), N-hydroxyethylpiperazine (HEP), N-methyldiethanolamine (MDEA), monoethanolamine (MEA), N-methylpiperazine (MP), piperazine, piperidine, 2-(2-tert-butylaminoethoxy)ethanol (TBEE), triethanolamine (TEA), triisopropanolamine (TIA), tris, 2-(2-aminoethoxy)ethanol, 2-(2-tert-butylaminopropoxy)ethanol, 2-(2-tert-amylaminoethoxy)ethanol, 2-(2-isopropylaminopropoxy)ethanol, 2-(2-(l -methyl- l-ethylpropylamino)ethoxy)ethanol, and mixtures thereof. In some embodiments, the amine compound is selected from AMP, MEA, MDEA, TIA, and mixtures thereof, and in one preferred embodiment the solution comprises the amine compound MDEA. Further, in the embodiments of the methods employing an amine compound in solution, the suitable conditions can comprise an amine compound concentration of from about 1 M to about 10 M, from about 2 M to about 8 M, from about 2.5 M to about 6.5 M, from about 3 M to about 5 M, at least about 2 M, at least about 3 M, at least about 4.2 M, or at least about 5 M.

[0158] Elevated temperatures are typically present when the method employs a solution comprising an amine compound are used to remove carbon dioxide from a gas stream. Thus, in some embodiments the method is carried out wherein the suitable conditions comprise a solution comprising an amine compound *(e.g.,* MDEA) and a temperature of from about 40°C to about 110°C, from about 40°C to about 90°C, from about 40°C to about 80°, from about 40°C to about 70°C, or from about 40°C to about 60°C.

[0159] Another known process for capturing CO_2 from a gas stream *(e.g.,* flue gas) uses a solution containing a high concentration of ammonia. Due to the high volatility of ammonia vapor the process is typically run at relatively low temperatures, and is sometimes referred to as a "chilled ammonia" process. Methods and conditions of the chilled ammonia process for CO_2 capture from a flue gas stream are described in *e.g.,* U.S. Pat. No. 7,641,717 B2, and U.S. Pat. Publ. No. 2009/0155889A1, each which is hereby incorporated by reference herein. Accordingly, in some embodiments of the

methods of removing carbon dioxide disclosed herein, a solution containing ammonia is used to facilitate carbon dioxide absorption from the gas streams. Such ammonia solutions can be used under suitable conditions comprising an ammonia concentration of about 1 M to about 8 M, from about 2 M to about 7 M, from about 3 M to about 6 M, at least about 1 M, at least about 2 M, at least about 3 M, at least about 4 M, or at least about 5 M, or at least about 5.6 M. Further in some embodiments of the methods, the solution comprising ammonia can be used at chilled temperatures *(e.g.,* for absorption) and/or elevated temperatures *(e.g.,* for desorption of carbon dioxide). Accordingly, in some embodiments, the method using a solution comprising ammonia can be carried out wherein the suitable conditions comprise a solution temperature of from about 0°C to about 20°C, from about 0°C to about 10°C, from about 5°C to about 15°C, from about 8°C to about 12°C, less than about 15°C, or less than about 10°C.

[0160] Some processes for CO_2 capture from a gas stream use contact with a solution comprising elevated concentration of carbonate ions (CO_3^{2-}). Typically, the carbonate ion is provided in the solution in the form of potassium carbonate (K_2CO_3) or sodium carbonate (Na_2CO_3). In such embodiments, the stability and activity of the carbonic anhydrase in the presence of carbonate ions is an important functional characteristic. Accordingly, in some embodiments, the method of removing CO_2 from a gas stream can be carried out wherein the suitable conditions comprise a solution comprising carbonate ion at a concentration of about 0.1 M CO_3^{2-} to about 5 M CO_3^{2-}, from about 0.2 M CO_3^{2-} to about 4 M CO_3^{2-}, or from about 0.3 M CO_3^{2-} to about 3 M CO_3^{2-}, at least about 0.2 M Na_2CO_3, at least about 0.4 M Na_2CO_3, or at least about 1 M Na_2CO_3.

[0161] Generally, in the methods of the present disclosure, the solution comprises an aqueous solvent (water or aqueous co-solvent system) that may be pH buffered or unbuffered. Generally, the CO_2 absorption reaction via hydration of carbon dioxide can be carried out by the carbonic anhydrase polypeptides over a pH range of about pH 9 or above or at a pH of about pH 10 or above, usually in the range of from about 8 to about 12. During the course of both the hydration and the dehydration reactions, the pH of the reaction mixture may change. The pH of the reaction mixture may be maintained at a desired pH or within a desired pH range by the addition of an acid or a base during the course of the reaction. Alternatively, the pH may be controlled by using an aqueous solvent that comprises a buffer. Suitable buffers to maintain desired pH ranges are known in the art and include, for example, carbonate, HEPES, triethanolamine buffer, and the like. The ordinary artisan will recognize that other combinations of buffering and acid or base additions known in the art may also be used. In some embodiments, the methods can be carried out in a solution at a basic pH that thermodynamically and/or kinetically favors the solvation of CO_2 - *e.g.,* from about pH 8 to about pH 12. Accordingly, in some embodiments, the rate is determined at a pH of from about pH 8 to about pH 12, from about pH 9 to about pH 11.5, or from about pH 9.5 to pH 11. In other embodiments, release (dehydration) of captured carbon dioxide *(e.g.,* as bicarbonate) is carried out at a pH of about 9

or below, usually in the range of from about pH 5 to about pH 9, or about pH 6 to about pH 9. In some embodiments, the dehydration is carried out at a pH of about 8 or below, often in the range of from about pH 6 to about pH 8.

[0162] In some embodiments, the methods of removing carbon dioxide from a gas stream disclosed herein, the solution can comprise an aqueous co-solvent system. For example, certain co-solvents or compounds can be added to the aqueous solution to reduce their degradative or corrosive properties. In some embodiments of the method, the solution is an aqueous co-solvent system comprising a ratio of water to a co-solvent from about 95:5 (v/v) to about 5:95 (v/v), in some embodiments, from about 90:10 (v/v) to about 10:90 (v/v), in some embodiments, from about 80:20 to about 20:80 (v/v), in some embodiments, from about 70:30 (v/v) to about 30:70 (v/v), and in some embodiments, from about 60:40 (v/v) to about 40:60 (v/v). The solvent component of an aqueous co-solvent system may be miscible with the aqueous component, providing a single liquid phase, or may be partly miscible or immiscible with the aqueous component, providing two liquid phases. The co-solvent system may be pre-formed prior to addition to the reaction mixture, or it may be formed *in situ* in the reaction vessel.

[0163] Co-solvent systems used in the methods typically comprise a solvent or compound that thermodynamically and/or kinetically favors the solvation of CO_2 from a gas-solvent interface. In some embodiments, the co-solvent in the aqueous solution is an amine compound *(e.g.,* AMP, MDEA, MEA, TEA, and/or TIA). In some embodiments of the methods disclosed herein, the solution can comprise a mixture or blend of amine compounds, and/or other compounds that facilitate the absorption of CO_2 into the solution, *e.g.,* ammonia, carbonate ions, strong base *(e.g.,* NaOH), and/or compounds such as dimethyl ether of polyethylene glycol (PEG DME).

[0164] In some embodiments, the aqueous co-solvent systems can have water and one or more organic solvents. In general, an organic solvent component of an aqueous co-solvent system is selected such that it does not completely inactivate the carbonic anhydrase enzyme. Appropriate co-solvent systems can be readily identified by measuring the enzymatic activity of the specified engineered carbonic anhydrase enzyme in the candidate solvent system, utilizing an enzyme activity assay, such as those described herein.

6. EXAMPLES

[0165] Various features and embodiments of the disclosure are illustrated in the following representative examples, which are intended to be illustrative, and not limiting.

EXAMPLE 1: Preparation and Screening of Carbonic Anhydrase Polypeptides Based on Wild-type Gene from *Desulfovibrio vulgaris.*

[0166] This example illustrates designing and optimizing the wild-type carbonic anhydrase gene from *Desulfovibrio vulgaris,* as well as further optimization and functional screening of the gene to

generate engineered polypeptides having increased solvent and thermostability under conditions suitable for CO_2 absorption from gas into a capture solvent.

[0167] Gene acquisition, synthesis, cloning, and expression: The gene encoding a wild-type *Desulfovibrio vulgaris* carbonic anhydrase polypeptide of SEQ ID NO: 2 was codon-optimized for expression in *E. coli* as the nucleotide sequence of SEQ ID NO: 1. The codon-optimized gene of SEQ ID NO: 1 was synthesized using oligonucleotides, generally composed of 42 nucleotides, and cloned into the expression vector pCKl 10900 under the control of a *lac* promoter. This expression vector also contains the PI5a origin of replication and the chloramphenicol resistance gene. Resulting plasmids containing the codon-optimized were transformed into *E. coli* W3110 using standard methods. The transformed wild-type gene sequence of SEQ ID NO: 1 was confirmed by standard sequencing techniques and the resultant expression of carbonic anhydrase activity by the transformed cells confirmed by high throughput activity assays as described below.

[0168] Preparation of engineered carbonic anhydrase library: Using the codon-optimized wild-type gene of SEQ ID NO: 1 as the starting point, a library of engineered variant genes was synthesized that targeted every residue from position X2 to position X223 of SEQ ID NO: 2 with substitutions of all 19 amino acids. The resulting engineered carbonic anhydrases polypeptide sequences, specific amino acid differences, and relative level of improvement are listed in Tables 2A and 2B.

[0169] Cloning of engineered carbonic anhydrase genes: As with the codon-optimized wild-type gene of SEQ ID NO: 1, the library of engineered variant genes was cloned into vector pCKl 10900 and expressed in *E. coli* W3110. Antibiotic resistant transformants were selected and processed to identify those expressing a CA with improved thermostability. Cell selection, growth, induced expression of CA variant enzymes and collection of cell pellets were as described below.

[0170] Picking: Recombinant *E. coli* colonies carrying a gene encoding CA were picked using a Q-Bot® robotic colony picker (Genetix USA, Inc., Boston, MA) into 96-well shallow well microtiter plates containing in each well 180 μL LB Broth, 1% glucose and 30 μg/mL chloramphenicol (CAM). Cells were grown overnight at 37°C with shaking at 200 rpm. A 10 μL aliquot of this culture was then transferred into 96-deep well plates containing 390 μL TB broth and 30 μg/mL CAM. After incubation of the deep-well plates at 37°C with shaking at 250 rpm for 2-3 hrs, recombinant gene expression within the cultured cells was induced by addition of IPTG to a final concentration of 1 mM, followed by addition of $ZnS0_4$ to a final concentration of 0.5 mM. The plates were then incubated at 37°C with shaking at 250 rpm for 18 hrs.

[0171] Preparation of clear lysate for assay: Cells were pelleted by centrifugation (4000 RPM, 10 min, 4°C), resuspended in 200 μ lysis buffer and lysed by shaking at room temperature for 2 hours. The lysis buffer contained 25 mM HEPES buffer, pH 8, 1 mg/mL lysozyme, and 500 μg/mL polymixin B sulfate (PMBS) and 1 mM dithiothreitol (DTT). After sealing the plates with

aluminum/polypropylene laminate heat seal tape (Velocity 11, Menlo Park, CA, Cat# 06643-001), they were shaken vigorously for 2 hours at room temperature. Cell debris was pelleted by centrifugation (4000 RPM, 10 min., 4°C) and the clear supernatant assayed directly or stored at 4°C until use.

[0172] High-throughput screening for improved stability in amine solvent, MDEA: Screening of the polypeptides encoded by the variant genes for carbonic anhydrases with improved stability in high concentrations of an amine solvent, MDEA, was carried out using the assays as follows. After lysis, 25 μL of cleared *E. coli* lysate was added to 96-well Costar® shallow round bottom plate, followed by addition of 75 μL of amine solvent challenge buffer (4 M MDEA, pH 10; pH adjusted using CO_2 gas) using a Biomek NXp robotic instrument (Beckman Coulter, Fullerton, CA). The resulting challenge solution MDEA solvent concentration was 3 M. Challenge buffers with increased MDEA concentrations of 5.33 M and 6.66 M were used to generate 4 M and 5 M MDEA challenge solutions, which also were similarly adjusted to pH 10 with CO_2. The plates were heat-sealed with aluminum/polypropylene laminate heat seal tape (Velocity 11, Menlo Park, CA, Cat# 06643-001) at 175°C for 2.5 seconds. The challenge reactions were heated for 24 h at the challenge temperature (42°C, 50°C, or 55°C). Control reactions were maintained at 25°C for 24 h. After 24 h, the plates were centrifuged at 4°C for 10 min to clarify the reaction mixtures. Carbonic anhydrase activity after challenge was measured using a bicarbonate dehydration assay as follows: 10 μL of cleared reaction mixture was added to a 96-well NUNC™ polystyrene shallow flat bottom plate containing 190 μL of a solution of 0.3 M MDEA, pH 8 (pH adjusted with $C0_2$ gas), 200 mM $KHCO_3$, 400 μM phenolphthalein. The rate of the dehydration reaction was determined as the slope of absorbance change at 25°C (or 45°C) assay solution temperature monitored at 550 nm (phenolphthalein as indicator) over time (30 minutes) on a SpectraMax M2 reader (Molecular Devices, Sunnyvale, CA). Engineered carbonic anhydrase samples showing greater than 1.3-fold improvement in activity relative to the wild-type polypeptide of SEQ ID NO: 2 under the same challenge conditions (positive control) were retested in triplicate using the same conditions.As noted in Tables 2A and 2B, HTP screening of engineered carbonic anhydrase polypeptides for amine solvent stability and thermostability has been carried out using at least five different challenge conditions/assays. Assay 1: challenge for 24 h at 42°C in 3 M MDEA solution followed by dehydration activity assay at 25°C; Assay 2: challenge for 24 h at 50°C in 3 M MDEA solution followed by dehydration activity assay at 25°C; Assay 5: challenge for 24 h at 50°C in 4 M MDEA solution followed by dehydration activity assay at 45°C; Assay 6: challenge for 24 h at 50°C in 5M MDEA solution followed by dehydration activity assay at 25°C; and Assay 7: challenge for 24 h at 55°C in 5 M MDEA solution followed by dehydration activity assay at 25°C. More stringent challenge conditions having higher amine solvent concentrations, and/or temperature, and/or additional reaction components *(e.g.,* potential inhibiting impurities) are contemplated for screening further engineered carbonic anhydrase polypeptides having

higher levels of stability and/or tolerance to the challenge conditions.High-throughput screening for improved stability in ammonia solvent: Screening of the engineered carbonic anhydrase polypeptides for improved stability in high concentrations of ammonia, was carried out using essentially the same HTP assay as for MDEA amine solvent described above but with the following changes. After lysis, 25 µï of lysate was transferred into 96-well Costar® shallow round bottom plates containing 75 µï of ammonia challenge buffer (5.6 M NH_3 (10 wt%) loaded with 0.3 molar equivalents of $C0_2$ gas). The resulting challenge solution ammonia concentration was 4.2 M (7.5 wt%). The challenge solutions were heated for 24 h at the challenge temperature (30°C or 35°C). Control solutions were maintained at 25°C for 24 h. After 24 h under challenge conditions, carbonic anhydrase activity was measured using a bicarbonate dehydration assay as follows: 10 µï of challenge (or control) solution was transferred to 190 µï of buffer (100 mM HEPES buffer, pH 7; 200 mM $KHC0_3$, 400 µM phenolphthalein). The rate of the dehydration reaction was determined as the slope of absorbance change at 25°C assay solution temperature monitored at 550 nm (phenolphthalein is a color indicator) over time (20 minutes). Engineered carbonic anhydrase samples showing greater than 1.3-fold improvement in activity relative to the wild-type polypeptide of SEQ ID NO: 2 under the same challenge conditions (positive control) were retested in triplicate using the same conditions.

[0175] As noted in Tables 2A and 2B, HTP screening of engineered carbonic anhydrase polypeptides for ammonia solvent stability and thermostability has been carried out using at least two different challenge conditions/assays. Assay 3: challenge for 24 h at 30°C in 4.2 M NH_3 solution containing 0.3 molar equivalents of CO_2 (a = 0.3), followed by dehydration activity assay at 25°C; and Assay 4: challenge for 24 h at 35°C in 4.2 M NH_3 solution containing 0.3 molar equivalents of CO_2 (a = 0.3), followed by dehydration activity assay at 25°C. More stringent challenge conditions having higher ammonia solvent concentrations, and/or higher or lower temperatures, and/or additional reaction components *(e.g.,* potential inhibiting impurities) are contemplated for screening further engineered carbonic anhydrase polypeptides having higher levels of stability and/or tolerance to the challenge conditions.

[0176] Production of recombinant carbonic anhydrase shake-flask powder (SFP): A shake-flask procedure was used to generate recombinant carbonic anhydrase polypeptide powders used in secondary screening assays or in the carbon capture processes disclosed herein. Shake flask powder (SFP) includes approximately 30% total protein and accordingly provide a more purified preparation of an engineered enzyme as compared to the cell lysate. A single microbial colony of *E. coli* containing a plasmid encoding a CA of interest was inoculated into 50 mL Luria Bertani broth containing 30 µg/mL chloramphenicol and 1% glucose. Cells were grown overnight (at least 16 hrs) in an incubator at 30°C with shaking at 250 rpm. The culture was diluted into 250 mL 2XYT media containing 30 µg/mL chloramphenicol, in a 1 liter flask to an optical density at 600 nm (OD_{600}) of 0.2 and allowed to grow at 30°C. Expression of the CA gene was induced by addition of isopropyl β D-

thiogalactoside (IPTG) to a final concentration of 1 mM when the OD_{600} of the culture was 0.6 to 0.8. $ZnSO_4$ was then added to a final concentration of 0.5 mM and incubation was then continued overnight (at least 16 hrs). Cells were harvested by centrifugation (5000 rpm, 15 min, 4°C) and the supernatant discarded. The cell pellet was resuspended with an equal volume of cold (4°C) 25 mM HEPES buffer, pH 8, and passed through a homogenizer twice at 33.6 kpsi while maintained at 4°C. Cell debris was removed by centrifugation (9000 rpm, 45 min., 4°C). The clear lysate supernatant was collected and stored at -20°C. Lyophilization of frozen clear lysate provides a dry powder (shake flask powder) of recombinant carbonic anhydrase polypeptide.

[0177] Production of recombinant carbonic anhydrase downstream-processed (DSP) powder: DSP powders contains approximately 80% total protein and accordingly provide a more purified preparation of the engineered carbonic anhydrase as compared to the cell lysate. Larger-scale (-100 - 120 g) fermentation of the engineered carbonic anhydrase for production of DSP powders can be carried out as a short batch followed by a fed batch process according to standard bioprocess methods.

[0178] A single microbial colony of *E. coli* containing a plasmid with the recombinant carbonic anhydrase gene of interest was inoculated into 2 mL M9YE broth containing 30 μg/mL chloramphenicol and 1% glucose. Cells were grown overnight (at least 12 h) in an incubator at 37°C with shaking at 250 rpm. After overnight growth, 0.5 mL of this culture was diluted into 250 mL M9YE Broth containing 30 μg/mL chloramphenicol and 1% glucose in 1 liter flask and allowed to grow at 37°C with shaking at 250 rpm. When the OD_{600} of the culture is 0.5 to 1.0, the cells were removed from the incubator and either used immediately, or stored at 4°C.

[0179] Bench-scale fermentations were carried out at 30°C in an aerated, agitated 15 L fermentor using 6.0 L of growth medium consisting of: 0.88 g/L ammonium sulfate, 0.98 g/L of sodium citrate; 12.5 g/L of dipotassium hydrogen phosphate trihydrate, 6.25 g/L of potassium dihydrogen phosphate, 3.3 g/L of Tastone-154 yeast extract, 0.083 g/L ferric ammonium citrate, and 8.3 mL/L of a trace element solution containing 2 g/L of calcium chloride dihydrate, 2.2 g/L of zinc sulfate heptahydrate, 0.5 g/L manganese sulfate monohydrate, 1 g/L cuprous sulfate heptahydrate, 0.1 g/L ammonium molybdate tetrahydrate and 0.02 g/L sodium tetraborate. The vessel was sterilized at 121°C and 15 PSI for 30 minutes, and ZnSC^ was added to 0.5 mM post sterilization. The fermentor was inoculated with a late exponential culture of *E. coli* W3110 containing a plasmid encoding the CA gene of interest (grown in a shake flask as described above to a starting OD_{600} of 0.5 to 1.0. The fermentor was agitated at 250-1250 rpm and air was supplied to the fermentation vessel at 0.6-25 L/min to maintain a dissolved oxygen level of 50% saturation or greater. The pH of the culture was maintained at 7.0 by addition of 20% v/v ammonium hydroxide. Growth of the culture was maintained by addition of a feed solution containing 500 g/L Cerelose dextrose, 12 g/L ammonium chloride and 5.1 g/L magnesium sulfate heptahydrate. After the culture reached an OD_{600} of 70±10, expression of CA was induced by addition of isopropyl -P-D-thiogalactoside (IPTG) to a final concentration of 1

mM and fermentation is continued for another 18 hours. The culture was then chilled to 4°C and maintained at that temperature until harvested. Cells were collected by centrifugation at 5000 G for 40 minutes in a Sorval RC12BP centrifuge at 4°C. Harvested cells were used directly in the following downstream recovery process or they may be stored at 4°C or frozen at -80°C until such use.

[0180] The cell pellet was resuspended in 2 volumes of 25 mM triethanolamine (sulfate) buffer, pH 7.5 at 4°C to each volume of wet cell paste. The intracellular CA was released from the cells by passing the suspension through a homogenizer fitted with a two-stage homogenizing valve assembly using a pressure of 12000 psig. The cell homogenate was cooled to -20°C immediately after disruption. A solution of 11% w/v polyethyleneimine pH 7.2 was added to the lysate to a final concentration of 0.5% w/v. A solution of 1 M Na_2SO_4 was added to the lysate to a final concentration of 100 mM. The lysate was then stirred for 30 minutes. The resulting suspension was clarified by centrifugation at 5000G in a Sorval RC12BP centrifuge at 4°C for 30 minutes. The clear supernatant was decanted and concentrated ten-fold using a cellulose ultrafiltration membrane with a molecular weight cut off of 10 kD. The final concentrate was dispensed into shallow containers, frozen at -20°C and lyophilized to provide the DSP powder. The recombinant carbonic anhydrase DSP powder was stored at -80°C.

EXAMPLE 2: Acceleration of CO_2 Absorption by the Carbonic Anhydrase from *Desulfovibrio vulgaris* (SEQ ID NO: 2) in Presence of Various Amine Compounds and Carbonate Ions and Elevated Temperatures

[0181] This example illustrates the ability of the wild-type beta-class carbonic anhydrase from *Desulfovibrio vulgaris* (SEQ ID NO: 2) and the engineered carbonic anhydrase polypeptides identified from HTP screening to accelerate the absorption of CO_2 gas into solutions containing high concentrations of various amine compounds *(e.g.,* MDEA), or Na_2CO_3, as well as the amine compound MDEA at various elevated temperatures.

[0182] Stirred cell reactor apparatus: A stirred cell reactor (SCR) was used to measure the acceleration of CO_2 absorption rate in the presence of carbonic anhydrase polypeptides of the present disclosure. The SCR consists of a hermetically-sealed cylindrical reactor vessel in which a gas and a liquid phase are mixed while their interface remains flat resulting in a mass transfer rate that is well known. The SCR allows the gas pressure and the gas and liquid temperatures to be controlled and monitored over time.

[0183] SCR assay method: Carbonic anhydrase polypeptide shake-flask powder (DSP can also be used) and the CO_2 capture solution of interest *(e.g.,* 4.2 M MDEA) are added to the reactor vessel. In some assays, the CO_2 capture solution is pre-loaded with a specific mole ratio of CO_2 per amine compound or ammonia defined by the term a. Pre-loading of a solution with CO_2 is carried out by first adding unloaded capture solution to the vessel, pressurizing the vessel with pure CO_2 gas and

mixing the solution until the CO_2 pressure drops to a certain level. The difference between the highest pressure and lowest pressure is used (with the ideal gas law) to calculate a of the solution.

[0184] Following addition of enzyme and solution to the vessel, the pressure in the SCR is reduced until the boiling point is reached, and the system is allowed to equilibrate until the pressure and temperature no longer change. A reservoir containing $C0_2$ (pure or a mixture) is connected to the SCR and a connecting valve is opened briefly allowing $C0_2$ to enter the SCR. Typically, the valve is opened until there is a change in pressure of approximately 10 psi when pure $C0_2$ is used. After closing the connecting valve, the drop in pressure in the SCR, which corresponds to the capture of CO_2 in solution, is monitored over time along with the gas and liquid temperatures. A control assay without the enzyme is also carried out.

[0185] Calculation of rate acceleration: The slope of the logarithm of the pressure drop in the SCR over time is used to calculate the overall pseudo-first order rate constant (k_0v) according to Eq. 1.

$$\text{slope} = \frac{A\ln P_{\infty}}{At} = -\frac{RT\sqrt{A}}{V_G He_{CO2}}\sqrt{k_{0V}D_{co_2}}$$

(Eq. 1)

[0186] From k_0v, the second order rate constant, k_2, can then be calculated according to Eq. 2.

$$r_{CO2} = k_{ov}[C0_2], \quad k_{ov} = k_{2,Base}[\text{Base}] + k_{2CA}[\text{CA}]$$

(Eq. 2)

[0187] The acceleration provided by a carbonic anhydrase polypeptide, or $E_{Cat,x\ g/L,}$ is calculated by dividing the rate, k_0v measured with a specified amount (X g/L) of the carbonic anhydrase by the rate, kov measured without enzyme, according to Eq. 3.

$$\text{Acceleration} = E_{CatXg/L} = \frac{k_{ov\ with\ X\ g/L\ carbonic\ anhydrase}}{k_{ov\ without\ carbonic\ anhydrase}}$$

(Eq. 3)

[0188] Certain equations and physical constants are used in calculating k_0v. For Eq. 1 and Eq. 2 to be valid, the reaction must be operated in the pseudo first order regime, which requires the following conditions: Hatta number ("Ha") > 2, and E_∞/Ha > 5 (E_∞ = infinite enhancement factor). The Hatta number, Ha, and infinite enhancement factor, E_∞, are determined according to Eq. 4 and Eq. 5, respectively.

$$Ha = \frac{\sqrt{k_{ov} \cdot D_{CO_2}}}{k_L}$$

(Eq. 4)

$$E_{\infty} = \sqrt{\frac{D_{CO_2}}{D_{Base}}} + \sqrt{\frac{D_{Base}}{D_{CO_2}}} \cdot \frac{[\text{Base}] \cdot H_{CO_2}}{Z_{CO_2} \cdot P_{CO_2}}$$

(Eq. 5)

[0189] The physical constants used for SCR assays in solutions containing MDEA are summarized in Table 4.

TABLE 4

Gas volume	325 mL
Liquid volume	175 mL
Interfacial area	3.03 x 10^{-3} m^2
Gas temperature	The average gas temperature during the part of the experiment where the slope is taken.
Liquid temperature	The average liquid temperature during the part of the experiment where the slope is taken.
Vapor pressure (P_{vap})	Taken from the average of the first 10 pressure readings before the CO_2 valve is opened. Alternatively it can be calculated from: 133.3*EXP(20.386-5130/T(K)) assuming water is the only compound giving a vapor pressure.
Liquid side mass transfer coefficient (k_L)	determined experimentally to be 4.47 x 10^{-5} m/s (*see e.g.*, Versteeg et al, Chem. Eng. Sci., 1987, 42, 1103-1119 for procedure).
Diffusivity of CO_2 (D_{CO2})	Calculated as a function of liquid temperature and mass fraction of MDEA by the correlation given in Sandall et al, J. Chem. Eng. data 1989, 34, 385-391.
Diffusivity of MDEA (D_{MDEA})	Calculated as a function of liquid temperature by the correlation given in Snijder et al., J. Chem. And Engi. Data, 1993, 38, 475-480.
Henry constant of CO_2 (H_{CO2})	Calculated as a function of liquid temperature and mass fraction of MDEA by the correlation given in Sandall et al, J. Chem. Eng. data 1989, 34, 385-391.
Stoichiometric coefficient of CO_2 (Z_{CO2})	1 for the MDEA system.

[0190] Results

[0191] As shown in Table 5, a loading of 1 g/L shake flask powder of the naturally occurring beta class carbonic anhydrase polypeptide of SEQ ID NO: 2 was capable of accelerating the absorption of $C0_2$ by solutions containing a range of amine solvents with no pre-loading of $C0_2$ (a = 0) at

concentration ranges from 1 M up to 4.2 M. The observed amount of acceleration was greatest in the 1 M solutions and generally decreased with increasing amine concentration. However, even in 4.2 M MDEA, the acceleration relative to the rate without enzyme was 15.8.

Table 5

	Acceleration of CO_2 absorption			
[Amine] (M)	Acceleration ($k_{OV,cat@1\ g/L}/k_{OV,uncat}$)			
	MDEA	**AMP**	**TEA**	**TIA**
1	52.3	4.15	49.8	95.1
2	27.1	1.54	43.3	85.1
3	22.4	1.23	14.5	16.1
4.2	15.8			
MDEA – Methyldiethanolamine AMP - 2-amino-2-methyl-1-propanol TEA – Triethanolamine TIA – Triisopropanolamine				

[0192] As shown in Table 6, a loading of 1 g/L shake-flask powder of the naturally occurring carbonic anhydrase polypeptide of SEQ ID NO: 2 was capable of accelerating the absorption of CO2 by a solution at 25°C containing 1 M Na_2C0_3 (with no pre-loading of CO2). The initial level of acceleration was 142-fold increased relative to the control solution without the biocatalyst. The enzyme maintained a high level of acceleration at least 65-fold increased relative to no biocatalyst even after 7 days in the solution at 25°C.

Table 6

1 g/L SEQ ID NO: 2, 1 M Na_2CO_3, 25°C (no CO_2 pre-loaded)	
Time (h)	**Acceleration**
0	142
19.95	119
45.64	110
95	71
168	65

[0193] Further SCR assays of the naturally occurring carbonic anhydrase polypeptide of SEQ ID NO: 2 were carried out at 40°C in solutions pre-loaded with $C0_2$ (a = 0.1) and containing 0.5 g/L of the polypeptide and 2.0 to 4.2 M MDEA. The assay solutions were monitored for up to 49 h. As shown by the results listed in Table 7, only 0.5 g/L of the polypeptide of SEQ ID NO: 2 was capable of initially accelerating the absorption of CO2 in solutions at 40°C containing 2 M to 4.2 M MDEA from about 11-fold to about 3-fold relative to the control solution without biocatalyst added. Further

even after 16 h or more in the 2 M to 4.2 M MDEA solutions at 40°C, the polypeptide of SEQ ID NO: 2 was capable of still accelerating CO_2 absorption by at least 2-fold relative to the control solution.

Table 7

Assay Sample	Time (h)	k_{OV} (s^{-1})	Acceleration
2 M MDEA (no enzyme)	0	10.8	1
2 M MDEA + 0.5 g/L enzyme	0	116.9	10.9
2 M MDEA + 0.5 g/L enzyme	20.4	30.2	2.8
2 M MDEA + 0.5 g/L enzyme	49.2	14.5	1.3
2.5 M MDEA (no enzyme)	0	12.6	1
2.5 M MDEA + 0.5 g/L enzyme	0	112.6	9.0
2.5 M MDEA + 0.5 g/L enzyme	21.1	38.4	3.1
2.5 M MDEA + 0.5 g/L enzyme	48.25	22.2	1.8
3 M MDEA (no enzyme)	0	15.0	1
3 M MDEA + 0.5 g/L enzyme	0	103.5	6.9
3 M MDEA + 0.5 g/L enzyme	19.92	37.5	2.5
3 M MDEA + 0.5 g/L enzyme	44.33	20.3	1.4
4.2 M MDEA (no enzyme)	0.0	12.4	1
4.2 M MDEA + 0.5 g/L enzyme	0.0	40.7	3.3
4.2 M MDEA + 0.5 g/L enzyme	16.5	28.2	2.3
4.2 M MDEA + 0.5 g/L enzyme	47.0	16.6	1.3

EXAMPLE 3: Acceleration of CO_2 Absorption by Engineered Carbonic Anhydrase Polypeptides in the Presence of MDEA in Solution at Elevated Temperatures

[0194] This example illustrates the ability of engineered carbonic anhydrase polypeptides identified from HTP screening to accelerate the absorption of CO_2 gas into amine solvent (MDEA) solutions at elevated temperatures.

[0195] Assays measuring rate of CO_2 hydration catalyzed by engineered carbonic anhydrase polypeptides of SEQ ID NO: 6, 16, 26, 30, 42, 84, and 186 (and wild-type of SEQ ID NO: 2) in MDEA solvent at 40°C and 50°C were carried out using the SCR and methods as described above in Example 2. As shown in Table 8, the relative improvement in stability in MDEA solvent exhibited by the polypeptides was determined as fold-improvement in residual activity at various time points and also as half-life ($t_{1/2}$) of $C0_2$ hydration activity.

Table 8

Polypeptide	$t_{½}$	Fold Improvement relative to SEQ ID NO: 2

SEQ ID NO:	(ho	$t_{½}$	Activity at Oh	Residual Activity at 24h	Residual Activity at 48h	Residual Activity at 72h
0.5 g/L polypeptide, 3 M MDEA, T = 40°C, <x=0.1						
2	20.0	1	1	1	1	1
6	67.8	3.4	2.18	2.2	5.0	4.8
16	69.9	3.5	1.60	2.6	5.4	4.9
30	30.7	1.5	1.45	1.7	2.6	1.8
42	53.0	2.7	1.78	2.2	4.2	4.1
84	18.2	0.9	1.74	1.0	1.2	0.7
186	24.7	1.2	1.68	1.5	2.1	1.5
0.5 g/L polypeptide, 3 M MDEA, T = 40°C, <x=0.02						
2	32.0	1	1	1	1	1
6	132.1	4.1	1.20	2.0	2.7	2.8
16	153.7	4.8	1.02	2.0	2.7	2.6
30	82.8	2.6	1.03	1.8	2.7	2.6
42	88.6	2.8	1.13	1.4	2.2	2.3
84	27.6	0.9	1.10	0.9	1.0	0.8
186	45.5	1.4	1.06	1.2	1.6	1.1
0.5 g/L polypeptide, 3 M MDEA, T = 50°C, <x=0.02						
2	0.13	1.0				
16	9.0	71.0				
26	9.8	77.0				

[0196] The engineered carbonic anhydrase polypeptides of SEQ ID NO: 6, 16, 30, 42, and 186 exhibited 1.2-fold to 4.8-fold increased stability in a 3 M MDEA solution at the elevated temperature of 40°C (with a CO_2 loading a = 0.02 or 0.1) when measured as $ti_{/2}$ for CO_2 absorption activity relative to the wild-type polypeptide of SEQ ID NO: 2. At the further elevated temperature of 50°C (with a CO_2 loading of a = 0.02) the engineered polypeptides of SEQ ID NO: 16 and 26 exhibited over 70-fold increased $ti_{/2}$ for CO_2 absorption activity in a 3 M MDEA solution relative to the wild type polypeptide of SEQ ID NO: 2.

[0197] The engineered carbonic anhydrase polypeptides of SEQ ID NO: 6, 16, 30, 42, and 186, continued to maintain their improved stability even at 48 h at 40°C. Additionally, in the case of the assays at the higher CO_2 loadings (a = 0.1), the stability increased significantly relative to that of the wild-type polypeptide of SEQ ID NO: 2 - *e.g.,* for SEQ ID NO: 16 increased from 1.6-fold to 5.4-fold greater than SEQ ID NO: 2.

EXAMPLE 4: Acceleration of CO_2 Absorption by the Carbonic Anhydrase from *Desulfovibrio vulgaris* (SEQ ID NO: 2) in Presence of Ammonia in Solution at Chilled Temperatures

[0198] This example illustrates the ability of the beta-class carbonic anhydrase from *Desulfovibrio vulgaris* (SEQ ID NO: 2) to accelerate the absorption of CO_2 gas into a chilled ammonia solution.

[0199] Apparatus and assay method

[0200] To a Parr Series 5100 low pressure reactor system fitted with a mass flow meter, a digital pressure gauge, a septum-capped addition/sampling port, a thermal well, a cooling loop (used as baffles/agitator shaft support) and a 450 mL glass jacketed cylinder was added water and the water degassed via vacuum at room temperature for ~20-40 minutes (until no bubble formation was observed). The cylinder was detached under a gentle nitrogen flow and 30 wt% NH_3 solution was added to make up the desired NH_3 solution with a final volume of ~250 mL *(e.g.,* 250 mL of 10wt% NH_3 solution = 166 mL of water and 83 mL of 30wt% NH_3). The 450 mL glass jacketed cylinder with the NH_3 solution was reattached to the reactor under a nitrogen atmosphere and the internal temperature was adjusted to the desired level via an external heat exchanger/circulator.

[0201] The turbine propeller was positioned on the stirrer shaft such that it was slightly above the liquid level and was used to mix the gas phase. An egg-shaped stir bar was placed in the cylinder and was used to stir the liquid phase via an external stir plate situated underneath the cylinder. Typically, the gas phase was stirred at 1800-2000 rpm and the liquid phase was stirred at 900-1200 rpm (fastest rate such that the surface of the liquid remained relatively flat/ripple-free). The internal temperature of the gas phase, the internal temperature of the liquid phase, the internal gas phase pressure, the agitation rates and the jacket temperature were recorded via a data logger.

[0202] After the internal temperature and pressure had equilibrated/stabilized, CO_2 gas was introduced through the mass flow meter until the desired initial loading of CO_2 was obtained. Loading was denoted as "a" which corresponds to the mole ratio of CO_2 to NH_3 *(e.g., a* = 0.3 means 3 moles of CO_2 per 10 moles of NH_3). Generally, depending on process optimization in an industrial scale process for CO_2 capture using chilled ammonia solution it is contemplated that the solution will enter the flue gas absorber at a relatively "lean" loading, of about a = 0.1 - 0.3 and after absorbing CO_2 will leave the absorber at a "rich" loading, dependent on equilibrium, of about a = 0.5 - 0.7.

[0203] Biocatalyst was introduced as an aqueous solution through the addition port. For control reaction, no additional solution was introduced. Then, for both sample and control reactions, a quick burst of CO_2 was added to the reactor vessel such that the partial pressure of CO_2 in the reactor was 5-15 psig. The vessel then was sealed. The subsequent decrease in the partial pressure of CO_2 in the reactor over time was recorded. The kinetic parameters were determined via analysis of the pressure versus time data under the prescribed reactor conditions. The composition of the solution in the reactor could also be monitored via samplings through the addition port. The acceleration in the rate of CO_2 absorption was calculated as described in Example 2.

[0204] Results

[0205] A set of assays were carried out at 10°C in a solution containing 5.6 M NH_3 with and without 2 g/L of the naturally occurring beta class carbonic anhydrase of SEQ ID NO: 2, with the CO_2 loading of the solution varied from a=0.30 to a=0.62.

[0206] As shown in Table 9, the observed rate constants, k_{ov}, with and without enzyme decreased with increased $C0_2$ loading in the solution *(i.e.,* increasing a), but k_{ov} increased as the $C0_2$ partial pressure in the gas phase decreased.

Table 9

	k_{ov} (s^{-1})			
	CO_2 partial pressure drop (atm)			
Sample Loading	**0.2 → 0.15**	**0.15 → 0.10**	**0.10 → 0.05**	**0.05 → 0.02**
α = 0.30 + enzyme	326	414	617	707
α = 0.30 control	110	137	179	193
α = 0.36 + enzyme	165	223	342	537
α = 0.36 control	33.0	43.3	58.8	89.8
α = 0.41 + enzyme	152	210	336	517
α = 0.41 control	14.4	17.4	23.6	32.4
α = 0.47 + enzyme	95	131	223	432
α = 0.47 control	5.3	7.4	13.6	33.5
α = 0.53 + enzyme	48	58	96	201
α = 0.53 control	3.1	3.5	3.5	2.5
α = 0.62 + enzyme	61	64	73	72
α = 0.62 control	3.8	4.4	6.1	11.3

[0207] As shown in Table 10, the naturally occurring carbonic anhydrase polypeptide of SEQ ID NO: 2 exhibited significant CO_2 absorption acceleration in the chilled ammonia solution (5.6 M NH_3 at 10°C).

Table 10

CO_2 Loading (α)	**Enzyme acceleration of CO_2 absorption**
0.30	3.0
0.36	5.5
0.41	13
0.47	17
0.53	22.5

[0208] The amount of acceleration by the presence of the polypeptide of SEQ ID NO: 2 increased linearly from a value of about 3.0, at a loading of a = 0.30, up to about 22.5, at a loading of a=0.53. Above a = 0.53 the pseudo first order behavior of k_{ov} appeared to break down and the rate of acceleration could not be determined accurately.

[0209] Further assays were carried out at 10°C in a solution containing 5.6 M NH_3, a solution CO_2 loading of the solution of a=0.30-0.40 and 2 g/L of a recombinant carbonic anhydrase from Table 2A. The recombinant carbonic anhydrases polypeptides had amino acid sequences of SEQ ID NO: 6, 26,

32, 60, and 124, and included the following amino acid residue difference relative to SEQ ID NO 2: X I5R, X30R; X56S, X86A, and X I 19K. All of the assayed recombinant carbonic anhydrases polypeptides accelerated the CO_2 absorption by the 5.6 M NH_3 solution at 10°C equivalent to the acceleration exhibited by wild-type of SEQ ID NO: 2. In contrast, the wild-type carbonic anhydrases of SEQ ID NO: 270, 272, and 274, each of which has some amino acid sequence homology to SEQ ID NO: 2 exhibited no observable acceleration over baseline of the $C0_2$ absorption by the 5.6 M NH_3 solution at 10°C. Thus, wild-type carbonic anhydrase polypeptide from *D. vulgaris* of SEQ ID NO: 2, or one of the engineered carbonic anhydrase polypeptides comprising one or more of the amino acid differences X I5R, X30R; X56S, and X I 19K, is capable of significantly accelerating carbon dioxide absorption by a solution under "chilled ammonia" process conditions of 5.6 M NH_3, a=0. 3-0.4, 2 g/L polypeptide, and T = 10°C.

[0210] All publications, patents, patent applications and other documents cited in this application are hereby incorporated by reference in their entireties for all purposes to the same extent as if each individual publication, patent, patent application or other document were individually indicated to be incorporated by reference for all purposes.

[0211] While various specific embodiments have been illustrated and described, it will be appreciated that various changes can be made without departing from the spirit and scope of the invention(s).

What is claimed is:

1. A recombinant carbonic anhydrase polypeptide having an improved enzyme property relative to a reference polypeptide of SEQ ID NO:2, said recombinant polypeptide comprising an amino acid sequence having at least 80% identity to SEQ ID NO:2 and an amino acid residue difference at one or more of the following positions relative to SEQ ID NO: 2: X2; X4; X15; X16; X22; X30; X31; X32; X34; X35; X37; X40; X42; X43; X47; X56; X60; X68; X70; X84; X86; X95; X96; X97; X I 19; X120; X121; X124; X131; X138; X139; X142; X143; X144; X145; X147; X148; X157; X159; X168; X170; X178; X200; X207; X213; X219; X221; X222; andX223.

2. The recombinant carbonic anhydrase polypeptide of claim 1 in which the amino acid sequence comprises one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2R; X2T; X4F; X4M; X15R; X16S; X22G; X30A; X30K; X30L; X30Q; X30R; X31P; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40Q; X40W; X42A; X43M; X43V; X47R; X56S; X60C; X60V; X68A; X68G; X70I; X84K; X84N; X84Q; X84R; X84S; X86A; X95V; X96A; X96C; X96E; X96K; X97F; X I 19K; X I 19L; X I 19M; X I 19T; X120R; X121H; X121K; X121L; X121Q; X121T; X121V; X121W; X124G; X124R; X131L; X138F; X138L; X138W; X139H; X139K; X139M; X139Q; X142L; X143M; X144A; X144L; X145C; X145F; X145W; X147E; X147F; X147G; X147T; X148A; X148T; X157A; X159H; X159R; X159V; X168E; X170F; X178G; X200R; X207E; X207N; X213E; X213Q; X219T; X221C; X222C; X223C; andX223Q.

3. The recombinant carbonic anhydrase polypeptide of claim 1 in which the improved property comprises increased stability in the presence of an amine compound and the amino acid sequence comprises amino acid residue differences at one or more of the following positions relative to SEQ ID NO: 2: X2; X4; X15; X16; X30; X31; X32; X34; X35; X37; X40; X42; X43; X47; X56; X60; X68; X70; X84; X86; X95; X96; X97; X I 19; X120; X121; X124; X131; X138; X139; X142; X144; X145; X147; X148; X159; X168; X170; X178; X200; X213; X219; X221; X222; and X223.

4. The recombinant carbonic anhydrase polypeptide of claim 3 in which the amino acid sequence comprises one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2R; X2T; X4F; X4M; X15R; X16S; X30A; X30K; X30L; X30Q; X30R; X31P; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40W; X42A; X43M; X43V; X47R; X56S; X60C; X68A; X68G; X70I; X84N; X84Q; X84R; X84S; X86A; X95V; X96A; X96C; X96E; X96K; X97F; X I 19K; X I 19L; X I 19M; X I 19T; X120R; X121H; X121K; X121L; X121Q; X121T; X121V; X121W; X124G; X124R; X131L; X138L; X139H; X139K; X139M; X139Q; X142L; X144A; X144L; X145C; X145F; X145W; X147E; X147F; X147G; X147T; X148A; X148T; X159H; X159R; X159V; X168E; X170F; X178G; X200R; X213E; X213Q; X219T; X221C; X222C; and X223C.

5. The recombinant carbonic anhydrase polypeptide of claim 4 in which the increased stability in the presence of an amine compound comprises at least 1.3-fold increased activity relative to the reference polypeptide of SEQ ID NO: 2 after 24 hours exposure to 4 M MDEA at 42°C.

6. The recombinant carbonic anhydrase polypeptide of claim 4 in which the increased stability in the presence of an amine compound comprises at least 1.5-fold increased activity relative to the reference polypeptide of SEQ ID NO: 2 after 24 hours exposure to 4 M MDEA at 50°C and the amino acid sequence comprises one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2T; X4F; X31P; X40L; X56S; X84Q; X119M; X120R; X121K; X121W; X131L; X139M; X147E; X147T; and X170F.

7. The recombinant carbonic anhydrase polypeptide of claim 4 in which the increased stability in the presence of an amine compound comprises at least 2-fold increased activity relative to the reference polypeptide of SEQ ID NO: 2 after 24 hours exposure to 4 M MDEA at 50°C and the amino acid sequence comprises one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2T; X56S; X84Q; X139M; X147E; and X147T.

8. The recombinant carbonic anhydrase polypeptide of claim 4 in which the increased stability in the presence of an amine compound comprises at least 3-fold increased activity relative to the reference polypeptide of SEQ ID NO: 2 after 24 hours exposure to 4 M MDEA at 50°C and an amino acid sequence comprising one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X56S; X84Q; X147E and X147T.

9. The recombinant carbonic anhydrase polypeptide of claim 3 in which the amino acid sequence comprises the amino acid difference X56S and one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2R; X2T; X4F; X4M; X15R; X16S; X30A; X30K; X30L; X30Q; X30R; X31P; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40W; X42A; X43M; X43V; X47R; X60C; X68A; X68G; X70I; X84N; X84Q; X84R; X84S; X86A; X95V; X96A; X96C; X96E; X96K; X97F; X119K; X119L; X119M; X119T; X120R; X121H; X121K; X121L; X121Q; X121T; X121V; X121W; X124G; X124R; X131L; X138L; X139H; X139K; X139M; X139Q; X142L; X144A; X144L; X145C; X145F; X145W; X147E; X147F; X147G; X147T; X148A; X148T; X159H; X159R; X159V; X168E; X170F; X178G; X200R; X213E; X213Q; X219T; X221C; X222C; andX223C.

10. The recombinant carbonic anhydrase polypeptide of claim 1 in which the improved property comprises increased stability in the presence of ammonia and the amino acid sequence comprises amino acid residue differences at one or more of the following positions relative to SEQ ID

NO: 2: X2; X4; X15; X22; X30; X32; X34; X35; X37; X40; X42; X47; X56; X60; X68; X70; X84; X86; X95; X96; X121; X124; X138; X143; X157; X170; X207; X219; X221; X222; andX223.

11. The recombinant carbonic anhydrase polypeptide of claim 10 in which the amino acid sequence comprises one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2R; X4F; X4M; X15R; X22G; X30A; X30K; X30L; X30Q; X30R; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40Q; X40W; X42A; X47R; X56S; X60C; X60V; X68A; X68G; X70I; X84K; X84Q; X86A; X95V; X96E; X121H; X121K; X121Q; X121T; X121W; X124G; X138F; X138W; X143M; X157A; X170F; X207E; X207N; X219T; X221C; X222C; and X223Q.

12. The recombinant carbonic anhydrase polypeptide of claim 10 in which the increased stability in the presence of ammonia comprises at least 1.3-fold increased activity relative to the reference polypeptide of SEQ ID NO: 2 after 24 hours exposure to ammonia at 30°C.

13. The recombinant carbonic anhydrase polypeptide of claim 10 in which the increased stability in the presence of ammonia comprises at least 1.3-fold increased activity relative to the reference polypeptide of SEQ ID NO: 2 after 24 hours exposure to ammonia at 35°C and the amino acid sequence comprises one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2R; X4F; X22G; X30A; X30L; X30Q; X3 OR; X32K; X32R; X34H; X35A; X35R; X37R; X40Q; X40W; X56S; X60C; X60V; X68A; X68G; X70I; X84K; X84Q; X86A; X95V; X121Q; X121T; X121W; X157A; X221C; X222C; andX223Q.

14. The recombinant carbonic anhydrase polypeptide of claim 10 in which the increased stability in the presence of ammonia comprises at least 3-fold increased activity relative to the reference polypeptide of SEQ ID NO: 2 after 24 hours exposure to ammonia at 35°C and the amino acid sequence comprises one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X4F; X22G; X30A; X30Q; X30R; X32K; X34H; X35A; X37R; X56S; X60C; X60V; X70I; X84Q; X121W; X221C; X222C; and X223Q.

15. The recombinant carbonic anhydrase polypeptide of claim 10 in which the amino acid sequence comprises one or more of the amino acid residue differences selected from X15R and X30R, and further comprises one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2R; X4F; X4M; X15R; X22G; X30A; X30K; X30L; X30Q; X30R; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40Q; X40W; X42A; X47R; X56S; X60C; X60V; X68A; X68G; X70I; X84K; X84Q; X95V; X96E; X121H; X121K; X121Q; X121T; X121W; X124G; X138F; X138W; X143M; X157A; X170F; X207E; X207N; X219T; X221C; X222C; andX223Q.

16. The recombinant carbonic anhydrase polypeptide of any one of claims 1 to 15 in which the amino acid sequence further comprises one or more of the following amino acid residue

differences relative to SEQ ID NO: 2: X2R; X2T; X4F; X4M; X I5R; X I6S; X30A; X30K; X30L; X30Q; X30R; X31P; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40W; X42A; X43M; X43V; X47R; X56S; X60C; X68A; X68G; X70I; X84N; X84Q; X84R; X84S; X86A; X95V; X96A; X96C; X96E; X96K; X97F; X I 19K; X I 19L; X I 19M; X I 19T; X120R; X121H; X121K; X121L; X121Q; X121T; X121V; X121W; X124G; X124R; X131L; X138L; X139H; X139K; X139M; X139Q; X142L; X144A; X144L; X145C; X145F; X145W; X147E; X147F; X147G; X147T; X148A; X148T; X159H; X159R; X159V; X168E; X170F; X178G; X200R; X213E; X213Q; X219T; X221C; X222C; andX223C.

17. The recombinant carbonic anhydrase polypeptide of claim 1 in which the amino acid sequence comprises at least one amino acid residue difference selected from each of at least two of the following seven sets of amino acid residue differences:

(a) X2R; X2T; X4F;

(b) X121H; X121K; X121L; X121Q; X121T; X121V; X121W; X144A; X144L; X178G;

(c) X139H; X139K; X139M;

(d) X30A; X30L; X30Q; X30R; X40L; X40W; X68A; X96A; X96C; X96E; X96K; X I 19K; X I 19L; X I 19M; X I 19T; X120R;

(e) X35R; X124G; X147E; X147F; X147G; X147T; X159H; X159R;

(f) X31P;

(g) X56S; X84N; X84Q; X84S.

18. The recombinant carbonic anhydrase polypeptide of claim 17 in which the amino acid sequence comprises one amino acid residue difference selected from each of at least two, three, four, five, six, or all seven of the sets of amino acid residue differences.

19. The recombinant carbonic anhydrase polypeptide of any one of claims 1 to 18, wherein said improved property comprises at least 1.2-fold, at least 1.3-fold, at least 1.5-fold, at least 2-fold, at least 3-fold, at least 4-fold, at least 5-fold, at least 10-fold, or at least 25-fold increased activity of hydrating carbon dioxide or dehydrating bicarbonate under suitable conditions.

20. The recombinant carbonic anhydrase polypeptide of claim 19, wherein the suitable conditions comprise a loading of solution CO_2 of from about a = 0.005 to about a = 0.4, from about a = 0.01 to about a = 0.3, a = 0.015 to about a = 0.25, a = 0.02 to about a = 0.2, less than about a = 0.3, less than about a = 0.25, or less than about a = 0.2.

21. The recombinant carbonic anhydrase polypeptide of claim 19, wherein the suitable conditions comprise a carbonic anhydrase polypeptide concentration of from about 0.1 g/L to about

10 g/L, about 0.25 g/L to about 7.5 g/L, about 0.5 g/L to about 5 g/L, less than 10 g/L, less than about 5 g/L, or less than about 2.5 g/L.

22. The recombinant carbonic anhydrase polypeptide of claim 19, wherein said increased activity is determined after heating the recombinant carbonic anhydrase polypeptide and the reference polypeptide at a temperature of from about 30°C to 60°C for a period of time of about 60 minutes to about 1440 minutes.

23. The recombinant carbonic anhydrase polypeptide of claim 19, wherein the suitable conditions comprise a solution comprising an amine compound selected from: 2-(2-aminoethylamino)ethanol (AEE), 2-amino-2-hydroxymethyl-l,3-propanediol (AHPD), 2-amino-2-methyl- 1-propanol (AMP), diethanolamine (DEA), diisopropanolamine (DIPA), N-hydroxyethylpiperazine (HEP), N-methyldiethanolamine (MDEA), monoethanolamine (MEA), N-methylpiperazine (MP), piperazine, piperidine, 2-(2-tert-butylaminoethoxy)ethanol (TBEE), triethanolamine (TEA), triisopropanolamine (TIA), tris, 2-(2-aminoethoxy)ethanol, 2-(2-tert-butylaminopropoxy)ethanol, 2-(2-tert-amylaminoethoxy)ethanol, 2-(2-isopropylaminopropoxy)ethanol, 2-(2-(l -methyl- l-ethylpropylamino)ethoxy)ethanol, and mixtures thereof.

24. The recombinant carbonic anhydrase polypeptide of claim 19, wherein the suitable conditions comprise a solution temperature of from about 40°C to about 110°C, from about 40°C to about 90°C, from about 40°C to about 80°, from about 40°C to about 70°C, or from about 40°C to about 60°C.

25. The recombinant carbonic anhydrase polypeptide of claim 19, wherein the suitable conditions comprise an amine compound concentration of from about 1 M to about 10 M, from about 2 M to about 8 M, from about 2.5 M to about 6.5 M, from about 3 M to about 5 M, at least about 2 M, at least about 3 M, at least about 4.2 M, or at least about 5 M.

26. The recombinant carbonic anhydrase polypeptide of claim 19, wherein the suitable conditions comprise a solution comprising ammonia at a concentration of about 1 M to about 8 M, from about 2 M to about 7 M, from about 3 M to about 6 M, at least about 1 M, at least about 2 M, at least about 3 M, at least about 4 M, or at least about 5 M, or at least about 5.6 M..

27. The recombinant carbonic anhydrase polypeptide of claim 26, wherein the suitable conditions comprise a solution temperature of from about 0°C to about 20°C, from about 0°C to about 10°C, from about 5°C to about 15°C, from about 8°C to about 12°C, less than about 15°C, or less than about 10°C.

28. The recombinant carbonic anhydrase polypeptide of claim 19, wherein the suitable conditions comprise a solution comprising carbonate ion at a concentration of about 0.1 M CO_3^{2-} to about 5 M CO_3^{2-}, from about 0.2 M CO_3^{2-} to about 4 M CO_3^{2-}, or from about 0.3 M CO_3^{2-} to about 3 M CO_3^{2-}.

29. The recombinant carbonic anhydrase polypeptide of any one of claims 1 to 29 which comprises an amino acid sequence selected from the group consisting of SEQ ID NO: 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, 180, 182, 184, 186, 188, 190, 192, 194, 196, 198, 200, 202, 204, 206, 208, 210, 212, 214, 216, 218, 220, 222, 224, 226, 228, 230, 232, 234, 236, 238, 240, 242, 244, 246, 248, 250, 252, 254, 256, 258, 260, 262, 264, 266, and 268.

30. A beta-class carbonic anhydrase polypeptide capable of hydrating carbon dioxide in a solution comprising an amine compound or ammonia, wherein the polypeptide comprises an amino acid sequence at least about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% identical to any one of SEQ ID NO:2, 270, 272, 274, 276, or 278, and has an alanine residue at position X60 relative to SEQ ID NO: 2.

31. The beta-class carbonic anhydrase polypeptide of claim 30, wherein the amino acid sequence further comprises one or more of the following amino acid residue differences relative to SEQ ID NO: 2: X2R; X2T; X4F; X4M; X15R; X16S; X30A; X30K; X30L; X30Q; X30R; X31P; X32K; X32R; X34H; X35A; X35R; X37R; X40L; X40W; X42A; X43M; X43V; X47R; X56S; X60C; X68A; X68G; X70I; X84N; X84Q; X84R; X84S; X86A; X95V; X96A; X96C; X96E; X96K; X97F; X I 19K; X I 19L; X I 19M; X I 19T; X120R; X121H; X121K; X121L; X121Q; X121T; X121V; X121W; X124G; X124R; X131L; X138L; X139H; X139K; X139M; X139Q; X142L; X144A; X144L; X145C; X145F; X145W; X147E; X147F; X147G; X147T; X148A; X148T; X159H; X159R; X159V; X168E; X170F; X178G; X200R; X213E; X213Q; X219T; X221C; X222C; and X223C.

32. A polynucleotide encoding a recombinant carbonic anhydrase polypeptide of any one of claims 1 to 31.

33. The polynucleotide of claim 32 which comprises a nucleotide sequence selected from the group consisting of SEQ ID NO: 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135,

137, 139, 141, 143, 145, 147, 149, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177, 179, 181, 183, 185, 187, 189, 191, 193, 195, 197, 199, 201, 203, 205, 207, 209, 211, 213, 215, 217, 219, 221, 223, 225, 227, 229, 231, 233, 235, 237, 239, 241, 243, 245, 247, 249, 251, 253, 255, 257, 259, 261, 263, 265, 267, 269, 271, 273, 275, 277, 279, 281, 283, and 285.

34. An expression vector comprising the polynucleotide of claims 32 or 33 operably linked to control sequences capable of directing expression of the encoded polypeptide in a host cell.

35. The expression vector of claim 34, wherein the control sequence comprises a promoter.

36. The expression vector of claim 35, wherein the promoter comprises an *E. coli* promoter.

37. The expression vector of claim 35, wherein the control sequence comprises a secretion signal.

38. A host cell comprising the expression vector of any one of claims 34-37.

39. The host cell of claim 38, wherein the host cell species is *E. coli.*

40. The host cell of claim 38, wherein the host cell is a *Streptomyces spp.* or *Bacillus spp.* selected from *Streptomyces coelicolor* (DSM 40680) and *Streptomyces lividans* (DSM 40434), *B. amyloliquefaciens, B. licheniformis, B. megaterium, B. stearothermophilus,* and *B. subtilis.*

41. The host cell of claim 38, wherein the host cell is from a filamentous fungal organism selected from *Aspergillus niger, Aspergillus nidulans, Aspergillus awamori, Aspergillus oryzae, Aspergillus sojae, Aspergillus kawachi, Trichoderma reesei, Chrysosporium lucknowense, Myceliophthora thermophilia, Fusarium venenatum, Neurospora crassa, Humicola insolens, Humicola grisea, Penicillum verruculosum, Thielavia terrestris,* and teleomorphs, anamorphs, synonyms or taxonomic equivalents thereof.

42. A method of producing a carbonic anhydrase polypeptide comprising:

(a) transforming a host cell with an expression vector polynucleotide encoding a recombinant carbonic anhydrase polypeptide of any one of claims 1 to 31;

(b) culturing said transformed host cell under conditions whereby said recombinant carbonic anhydrase polypeptide is produced by said host cell; and

(c) recovering said recombinant carbonic anhydrase polypeptide from said host cells.

43. The method of claim 40, wherein said expression vector comprises a secretion signal.

44. A formulation comprising a carbonic anhydrase polypeptide of any one of claims 1 to 31, and a CO_2 absorbent solution comprising an amine compound or ammonia.

45. The formulation of claim 44, wherein the amine compound is selected from the group consisting of: 2-(2-aminoethylamino)ethanol (AEE), 2-amino-2-hydroxymethyl-l,3-propanediol (AHPD), 2-amino-2-methyl- 1-propanol (AMP), diethanolamine (DEA), diisopropanolamine (DIPA), N-hydroxyethylpiperazine (HEP), N-methyldiethanolamine (MDEA), monoethanolamine (MEA), N-methylpiperazine (MP), piperazine, piperidine, 2-(2-tert-butylaminoethoxy)ethanol (TBEE), triethanolamine (TEA), triisopropanolamine (TIA), tris, 2-(2-aminoethoxy)ethanol, 2-(2-tert-butylaminopropoxy)ethanol, 2-(2-tert-amylaminoethoxy)ethanol, 2-(2-isopropylaminopropoxy)ethanol, 2-(2-(l -methyl- l-ethylpropylamino)ethoxy)ethanol, and mixtures thereof.

46. The formulation of claim 44, wherein the recombinant carbonic anhydrase polypeptide is immobilized on a surface.

47. A method for removing carbon dioxide from a gas stream comprising the step of contacting the gas stream with a solution comprising a carbonic anhydrase polypeptide of any one of claims 1 to 31 under suitable conditions, whereby the solution absorbs at least a portion of the carbon dioxide from the gas stream.

48. The method of claim 47, wherein the solution comprises an amine compound selected from: 2-(2-aminoethylamino)ethanol (AEE), 2-amino-2-hydroxymethyl- 1,3-propanediol (AHPD), 2-amino-2-methyl- 1-propanol (AMP), diethanolamine (DEA), diisopropanolamine (DIPA), N-hydroxyethylpiperazine (HEP), N-methyldiethanolamine (MDEA), monoethanolamine (MEA), N-methylpiperazine (MP), piperazine, piperidine, 2-(2-tert-butylaminoethoxy)ethanol (TBEE), triethanolamine (TEA), triisopropanolamine (TIA), tris, 2-(2-aminoethoxy)ethanol, 2-(2-tert-butylaminopropoxy)ethanol, 2-(2-tert-amylaminoethoxy)ethanol, 2-(2-isopropylaminopropoxy)ethanol, 2-(2-(l -methyl- l-ethylpropylamino)ethoxy)ethanol, and mixtures thereof.

49. The method of claim 48, wherein the amine compound is selected from: AMP, DEA, MEA, MDEA, TEA, TIA, and mixtures thereof.

50. The method of claim 48, wherein the suitable conditions comprise a solution temperature of from about 40°C to about 110°C, from about 40°C to about 90°, from about 40°C to about 80°, or from about 40°C to about 70°C.

51. The method of claim 48, wherein the suitable conditions comprise an amine compound concentration of from about 1 M to about 10 M, from about 2 M to about 8 M, from about 2.5 M to about 6.5 M, from about 3 M to about 5 M, at least about 2 M, at least about 3 M, at least about 4.2 M, or at least about 5 M.

52. The method of claim 48, wherein the suitable conditions comprise a loading of solution CO_2 of from about a = 0 to about a = 0.6, from about a = 0.01 to about a = 0.5, a = 0.02 to about a = 0.4, a = 0.05 to about a = 0.3, a = 0.1 to about a = 0.4, a = 0.2 to about a = 0.3, less than about a = 0.4, less than about a = 0.3, or less than about a = 0.2.

53. The method of claim 48, wherein the suitable conditions comprise a lean loading of solution $C0_2$ about a = 0 to about a = 0.02 and a rich loading of solution $C0_2$ of from about a = 0.2 to about a = 0.5.

54. The method of claim 47, wherein the solution comprises ammonia.

55. The method of claim 54, wherein said suitable conditions comprise a solution temperature of from about 0°C to about 20°C, from about 0°C to about 10°C, from about 5°C to about 15°C, from about 8°C to about 12°C, or at about 10°C.

56. The method of claim 54, wherein said suitable conditions comprise an ammonia concentration of about 1 M to about 8 M, from about 2 M to about 7 M, from about 3 M to about 6 M, at least about 1 M, at least about 2 M, at least about 3 M, at least about 4 M, or at least about 5 M, or at least about 5.6 M.

57. The method of claim 54, wherein the suitable conditions comprise a loading of solution CO_2 of from about a = 0 to about a = 0.7, from about a = 0.1 to about a = 0.7, a = 0.1 to about a = 0.5, a = 0.1 to about a = 0.3, a = 0.4 to about a = 0.7, a = 0.5 to about a = 0.7, less than about a = 0.7, less than about a = 0.5, or less than about a = 0.3.

58. The method of claim 54, wherein the suitable conditions comprise a lean loading of solution CO_2 about a = 0.1 to about a = 0.3 and a rich loading of solution CO_2 of from about a = 0.5 to about a = 0.7.

59. The method of claim 47, wherein the solution comprises carbonate ion.

60. The method of claim 59, wherein the suitable conditions comprise from about 0.1 M $C0_3^{2-}$ to about 5 M $C0_3^{2-}$, from about 0.2 M $C0_3^{2-}$ to about 4 M $C0_3^{2-}$, or from about 0.3 M $C0_3^{2-}$ to about 3 M $C0_3^{2-}$.

61. The method of claim 47, wherein the suitable conditions comprise a carbonic anhydrase polypeptide concentration of from about 0.1 g/L to about 10 g/L, about 0.25 g/L to about 7.5 g/L, about 0.5 g/L to about 5 g/L, less than 10 g/L, less than about 5 g/L, or less than about 2.5 g/L.

62. The method of claim 47, wherein the carbonic anhydrase polypeptide is immobilized on a surface.

63. The method of claim 47, wherein the method further comprises isolating the immobilized carbonic anhydrase from the solution.

64. The method of claim 47, wherein the method further comprises exposing the solution comprising the carbonic anhydrase polypeptide and absorbed carbon dioxide to suitable conditions for desorbing the carbon dioxide from the solution.

65. The method of claim 64, wherein the suitable conditions for desorbing the carbon dioxide from the solution comprise heating the solution.

66. The method of claim 64, wherein the suitable conditions for desorbing the carbon dioxide from the solution comprise exposing the solution to low pressure or a vacuum.

67. The method of claim 64, wherein the suitable conditions for desorbing the carbon dioxide from the solution comprise adding a recombinant carbonic anhydrase polypeptide of any one of claims 1-33.

68. The method of claim 47, wherein the method further comprises isolating the solution comprising hydrated carbon dioxide and contacting the isolated solution with a carbonic anhydrase polypeptide of any one of claims 1-31, thereby converting hydrated carbon dioxide to carbon dioxide gas and water.