Table 1 Distribution of 280 000 chemical compounds over the 32 point groups. Note somewhat different frequencies for inorganic (I) and organic (O) compounds. (data collected by G. Johnson and published in [7]).

	I	O		I	O
1	0.67%	1.24%	422	0.40%	0.48%
$\bar{1}$	13.87	19.18	4mm	0.30	0.09
2	2.21	6.70	$\bar{4}$2m	0.82	0.34
M	1.30	1.46	4/mmm	4.53	0.69
2/m	34.63	44.81	6	0.41	0.22
222	3.56	10.13	$\bar{6}$	0.07	0.01
mm2	3.32	3.31	6/m	0.82	0.17
mmm	12.07	784	622	0.24	0.05
3	0.36	0.32	6mm	0.45	0.03
$\bar{3}$	1.21	0.58	$\bar{6}$m2	0.41	0.02
32	0.54	0.22	6/mmm	2.82	0.05
3m	0.74	0.22	23	0.44	0.09
$\bar{3}$m	3.18	0.25	m3	0.84	0.15
4	0.19	0.25	432	0.13	0.01
$\bar{4}$	0.25	0.18	$\bar{4}$3m	1.42	0.11
4/m	1.17	0.67	m3m	6.66	0.12

this book. Among the advantages are (i) the explicit calculation of the optimized quantity of interest (e.g., the energy), (ii) unbiased search techniques for exploring the energy landscape can – unlike the previously mentioned approaches, assuming knowledge of material's chemistry and likely crystal structures – arrive at completely unexpected results and truly novel structures. For instance, who would guess (based on whatever chemical knowledge) that boron under pressure would assume a NaCl-type structure composed of B_2 and B_{12} clusters with partially ionic bonding between the two? Who would guess that, when compressed to 2 million atmospheres, sodium assumes a structure unknown for any other element and becomes a transparent dielectric? Nevertheless, this is exactly what happens [13, 14], and these phenomena were first predicted using optimization techniques and only then confirmed experimentally.

When considering crystal structure prediction as an optimization problem – i.e. the problem of finding the global minimum of the energy landscape, certain properties of this landscape need to be explored. First, the number of distinct points on the landscape can be estimated as:

$$C = \binom{V/\delta^3}{N} \prod_i \binom{N}{n_i} \tag{1}$$

where N is the number of atoms in the unit cell of volume V, δ is a relevant discretization parameter (for instance, 1 Å) and n_i is the number of atoms of i-th
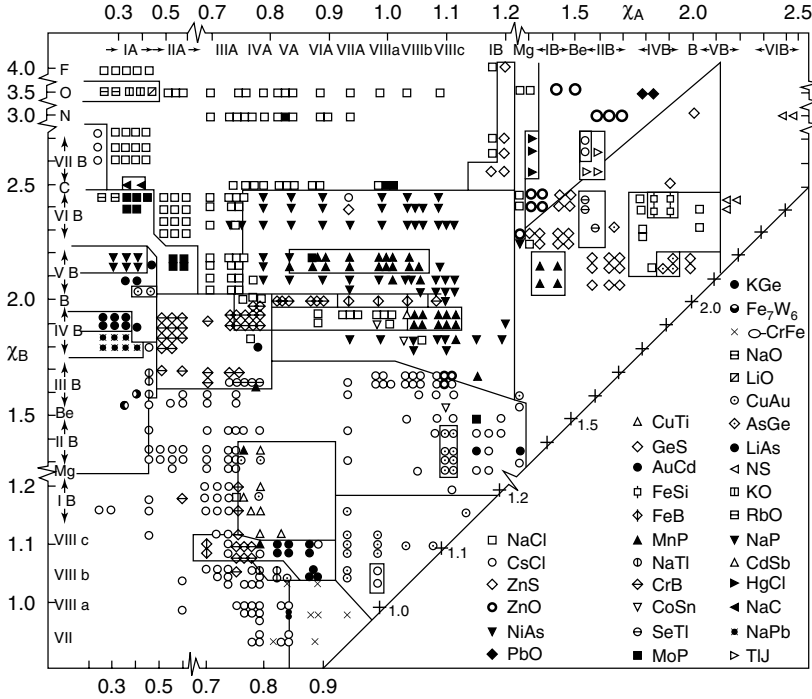
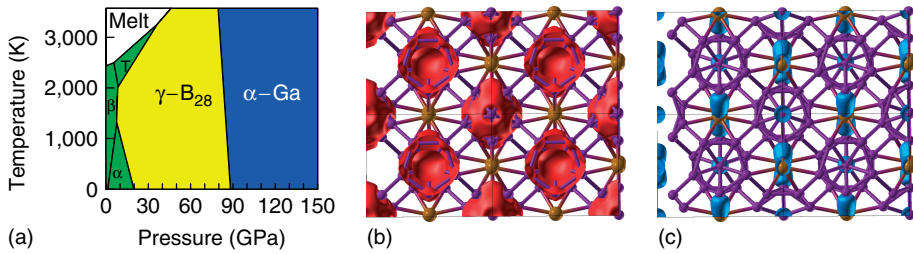**Figure 2** Pettifor's structure diagram for 574 AB compounds (from [10]).

Figure 3 Boron: (a) its schematic phase diagram (from [13]) and distribution of electrons corresponding (b) bottom and (c) top of the valence band in γ-B$_{28}$ [15].

type in the unit cell. C is astronomically large (roughly, $\sim 10^N$ if one uses $\delta = 1\,\text{Å}$ and typical atomic volume of $10\,\text{Å}^3$).

It is useful to consider the dimensionality of the energy landscape:

$$d = 3N + 3 \tag{2}$$

where $3N$-3 degrees of freedom are the atomic positions, and the remaining six dimensions are lattice parameters. For a system with 100 atoms in the unit cell, the landscape is 303-dimensional!